Phase Information in Robust Control (PIRC)

Final Report

David C. Hyland

Principal Investigator

Harris Corporation

MS 19/4848

Melbourne, FL 32902
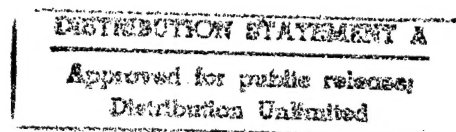
For:

Air Force Office of Scientific Research (AFOSR)

Bolling Air Force Base

Washington, DC 20332

Attention:

Dr. Spencer Wu

January 1995

19961209 084

DISCLAIMER NOTICE

UNCLASSIFIED

Technical Report
distributed by

DEFENSE
TECHNICAL
INFORMATION
CENTER

DTIC

UNCLASSIFIED

THIS DOCUMENT IS BEST
QUALITY AVAILABLE. THE
COPY FURNISHED TO DTIC
CONTAINED A SIGNIFICANT
NUMBER OF PAGES WHICH DO
NOT REPRODUCE LEGIBLY.

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE 24 Feb. 95	3. REPORT TYPE AND DATES COVERED 1 Jan. 92--31 Dec. 94

4. TITLE AND SUBTITLE
Phase Information in Robust Control (PIRC)

5. FUNDING NUMBERS
Contract #
F49620-92-C-0019

6. AUTHOR(S)
David C. Hyland and Emmanuel G. Collins

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Harris Corporation
Government Aerospace Systems Division
P.O. Box 94000
Melbourne, FL 32902-9400

8. PERFORMING ORGANIZATION REPORT NUMBER
1732-002

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
AFOSR/NA
Bolling AFB
Washington, DC 20332-6448

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

To design high performance, practically implementable control laws, it is important to have the appropriate tools for design and analysis. These tools should enable the following: (1) they should be based on robustness theory that is nonconversative with respect to the type of uncertainty being considered; (2) they should allow performance to be measured in a meaning full way; (3) they should yield controllers that are of sufficiently low order to be implemented on control processors with limited throughout capabilities; (4) they should be implemented via efficient numerical algorithms. The research cited in this final report has led to the further development of robustness theories and algorithms which include phase information regarding the uncertainty. In addition, this research has expanded the theory of optimal and suboptimal reduced-order control design and led to the development of new continuation algorithms for H_2 optimal reduced-order modeling and control based on the optimal projection equations. Finally, a new fixed-structure approach to complex structured singular value controller synthesis has been developed. The approach *a priori* constrains the order of the D-scales in the optimization process and can lead to much more robust controllers than standard D-K iteration and curve fitting approaches.

DTIC QUALITY INSPECTED 3

14. SUBJECT TERMS	15. NUMBER OF PAGES
Robust Control, Robustness Analysis, Model Reduction, Reduced-Order Control, Continuation Algorithms	354
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
unclassified	unclassified	unclassified	SAR

HARRIS

Marsha A. Carter
M/S: 22/4234, Rm 2225
Phone (407) 727-6177
FAX (407)-729-7157

October 21, 1996

TO: Pat Mawby

FROM: Marsha Carter

SUBJECT: Phase Information in Robust Contract (PIRC) Study -
            FINAL REPORT

REFERENCE: 1)   Contract No. F49620-92-C-0019
            2)   Telecon between P. Mawby and M. Carter
                  on 21 October 1996

As discussed during the referenced telecon, the missing pages from Appendix G of the subject final report cannot be located.

Should you have any questions, please contact the undersigned.

M. A. Carter
Senior Contract Administrator

Abstract

To design high performance, practically implementable control laws, it is important to have the appropriate tools for design and analysis. These tools should enable the following: 1.) they should be based on robustness theory that is nonconservative with respect to the type of uncertainty being considered; 2.) they should allow performance to be measured in a meaningful way; 3.) they should yield controllers that are of sufficiently low order to be implemented on control processors with limited throughput capabilities; 4.) they should be implemented via efficient numerical algorithms. The research cited in this final report has led to the further development of robustness theories and algorithms which include phase information regarding the uncertainty. In addition, this research has expanded the theory of optimal and suboptimal reduced-order control design and led to the development of new continuation algorithms for H_2 optimal reduced-order modeling and control based on the optimal projection equations. Finally, a new fixed-structure approach to complex structured singular value controller synthesis has been developed. This approach *a priori* constrains the order of the D-scales in the optimization process and can lead to much more robust controllers than standard D-K iteration and curve fitting approaches.

Table of Contents

1. Introduction

To design high performance, practically implementable control laws, it is important to have the appropriate tools for design and analysis. These tools should enable the following:

1. They should be based on robustness theory that is nonconservative with respect to the type of uncertainty being considered;

2. They should allow performance to be measured in a meaningful way;

3. They should yield controllers that are of sufficiently low order to be implemented on control processors with limited throughput capabilities;

4. They should be implemented via efficient numerical algorithms.

The ultimate aim of this research is to develop control design and analysis tools with the above characteristics.

From a theoretical perspective, to accomplish nonconservatism with respect to real, constant uncertainty or with respect to systems that are inherently stable (such as positive real systems), it is important to develop robustness theories that are not totally dependent on norms. This is because norm-based tests do not allow the inclusion of phase information regarding the uncertainty. These theories would be expected to deviate significantly from the popular but norm-based small gain tests.

Often times in the design of controllers for flexible structures, higher frequency modes are deleted from the model in order to enable the design of lower order controllers. The unmodeled modes are then accounted for as unstructured (i.e., magnitude bounded but arbitrary phase) uncertainty. If the unmodeled dynamics are actually fairy well known, an alternative is to include them in the control design model and design a reduced-order controller. The low-order controller can be designed by reducing the dimension of an optimal full-order controller or by direct design (i.e., directly optimizing some cost function). Hence, in this design process the structure of the controller is constrained a priori.

Another important fixed-structure problem which appears in robust control is a priori constraining the order of the D-scales in complex structured singular value (CSSV) controller synthesis. This robust design technique enables the design of controllers that are robust with respect to multiple block, unstructured uncertainty and also guarantee a certain measure of robust performance. However, current techniques for CSSV controller synthesis require the fitting of potentially very high

order D-scales with lower order approximations to avoid extremely high order controllers. This curve-fitting step can be very suboptimal and can even lead to a degradation of robust stability and performance in comparison with a standard H_∞ design. This highlights the importance of developing CSSV controller synthesis techniques that optimize the D-scales subject to a constraint on the D-scale order.

The above discussion motivates the objectives of PIRC. One objective was to extend majorant analysis to handle positive-real systems and specialize the basic theory to the case of collocated, decentralized, static rate feedback. The next objective was to extend the results to the case of collocated, dynamic rate feedback. In addition, we desired to compare the positive real majorant bounds with a totally norm based majorant bound and a performance bound obtained from complex structured singular value theory. Finally, we aimed to extend the results to the more realistic case in which the sensor and actuater dynamics are included in the plant model.

Like positive real majorant theory, Popov theory also enables the incorporation of phase information regarding the uncertainty. Our consultant, Dr. Wassim Haddad's first objective here was to extend Popov theory to handle bidirectional uncertainty analysis. We also desired to develop a special-case numerical algorithm to implement Popov robustness analysis. Next, we aimed to develop a more general algorithm and apply it to a realistic example. Due to our collaboration with Dr. Jonothan How at MIT, the analysis was to be performed using the Middeck Active Control Experiment (MACE).

The development of Popov robustness theory was largely motivated by the early work of Harris Corporation in Maximum Entropy control design, which has been shown empirically to nonconservatively execute the design of robust controllers for flexible structures with modal uncertainties. This research sought to develop a rigorous theoretical foundation for Maximum Entropy design and also to develop more efficient numerical algorithms for Maximum Entropy design.

An additional objective of this research was to develop continuation algorithms for optimal, reduced-order control design based on the optimal projection equations as opposed to the gradient expressions. Gradient-based methods directly optimize the controller parameters. To keep the number of controller parameters from becoming too large the controller is constrained to a minimal parameter basis. However, this constraint tends to introduce numerical ill-conditioning since the assumptions behind a minimal parameter basis are not always satisfied along the homotopy path or may be "poorly satisfied." The advantage of a gradient-based approach is that it easily enables the

development of "globally convergent homotopy algorithms," i.e., homotopy algorithms for which a well-conditioned homotopy map is guaranteed with probability one. An alternative which avoids the ill-conditioning due to the controller basis constraint is to develop a continuation algorithm based directly on the optimal projection equations. This class of algorithms does not currently fit into globally convergent homotopy theory, but for certain problems algorithms of this type can be implemented very efficiently and have exhibited good numerical robustness.

A final objective of this research was to develop a CSSV controller synthesis technique that constrains the order of the D-scales in the optimization process. This research has the potential to significantly impact a very important area of robust control design by developing a more reliable and optimal CSSV synthesis process. An enumeration of the research objectives is given below. .

Research Objectives

1. For the case of uncertain, strictly positive real plants controlled by positive real compensators, use majorant analysis to develop frequency domain performance bounds that are less conservative than previous majorant results. Then extend these results to the more realistic case of plants with sensor and actuater dynamics.

2. Extend recent work in Popov robustness theory, wherein the uncertainty is assumed to vary in only one direction (positive or negative), so that the uncertainty is allowed to vary in both directions.

3. Use Lyapunov functions to provide a more rigorous foundation for Maximum Entropy design.

4. Develop a continuation algorithm for Maximum Entropy design that can exhibit quadratic convergence properties along the continuation path.

5. Apply and compare both frequency domain and state-space versions of the Popov test to a benchmark problem. The state-space tests were to be applied via homotopy algorithms.

6. Develop a general algorithm for Popov analysis (with bi-directional uncertainty) and apply it to the Middeck Active Control Experiment (MACE) at MIT.

7. To better understand the relationship between the optimal projection equations for H_2 optimal reduced-order design and suboptimal controller reduction methods, extend optimal projection theory to the case in which the controller is not a priori assumed to be minimal (the standard assumption of optimal projection theory). Also, compare the projections used by the subopti-

mal methods that are able to produce a minimal realization of a nonminimal LQG compensator with the optimal projection.

8. To aid in the development of a rigorous initialization technique for continuation and homotopy algorithms for H_2 or H_2/H_∞ optimal design, develop a method for constructing nearly nonminimal LQG compensators.

9. As a step in developing a continuation algorithm for H_2 optimal reduced-order control design using the optimal projection equations, develop a continuation algorithm for the easier problem of H_2 optimal model reduction using the optimal projection equations.

10. Develop a continuation algorithm for H_2 optimal, reduced-order control design.

11. As a foundation for CSSV controller synthesis with fixed-order D-scales, use recent H_2/H_∞ theory to develop a CSSV analysis technique for constant D-scales and then fixed-order dynamic D-scales.

12. Develop a CSSV controller synthesis technique for constant D-scales.

2. Professional Personnel and Research Publications .

Personnel

This contract sponsored the research efforts of Dr. David C. Hyland, Dr. Emmanuel G. Collins, Jr., and Mr. Stephen Richter of Harris Corporation and Dr. Wassim M. Haddad of the Georgia Institute of Technology.

Journal Publications .

2.1 D. S. Bernstein, W. M. Haddad, D. C. Hyland, and F. Tyan, " Maximum Entropy-Type Lyapunov Functions for Robust Stability and Performance Analysis," *System and Control Letters*, Vol. 21, pp. 73-87, 1993. <Contained in Appendix A.>

2.2 E. G. Collins, Jr., L. D. Davis, and S. Richter, "Homotopy Algorithm for Maximum Entropy Design," *Journal of Guidance, Control, and Dynamics*, Vol. 17, No. 2, pp. 311-321, 3 March 1994. <Contained in Appendix B.>

2.3 W. M. Haddad and D. S. Bernstein, "The Multivariable Parabola Criterion for Robust Controller Synthesis: A Riccati Equation Approach," *Journal of Mathematical Systems, Estimation and Control,* to appear. <Contained in Appendix C.>

2.4 W. M. Haddad, E. G. Collins, Jr., and D. S. Bernstein, "Robust Stability Analysis Using the Small Gain, Circle, Positivity, and Popov Theorems: A Comparative Study," *IEEE Transactions on Control Systems Technology*, Vol. 1, pp. 290-293, Dec. 1993. <Contained in Appendix D.>

2.5 E. G. Collins, Jr., W. M. Haddad and L. D. Davis, "Riccati Equation Approaches for Robust Stability and Performance Analysis Using the Small Gain, Positivity, and Popov Theorems," *Journal of Guidance, Control, and Dynamics*, Vol. 17, No. 2, pp. 322-329, March 1994. <Contained in Appendix E.>

2.6 D. C. Hyland, E. G. Collins, Jr., W. M. Haddad, and V. S. Chellaboina, "Frequency Domain Performance Bounds for Uncertain Positive Real Plants Controlled By Strictly Positive Real Compensators," submitted to *International Journal of Dynamics and Control.* <Contained in Appendix F.>

2.7 J. P. How, E. G. Collins, Jr., and W. M. Haddad, "Optimal Popov Controller Analysis and

Synthesis for Systems with Real Parameter Uncertainty," submitted to *IEEE Transactions on Control Systems Technology.* <Contained in Appendix G.>

2.8 E. G. Collins, Jr., W. M. Haddad, and S. S. Ying, "Generalized Fixed-Structure Optimality Conditions for H_2 Optimal Control," submitted to *SIAM J. Control and Optimization.*<Contained in Appendix H.>

2.9 E. G. Collins, Jr., W. M. Haddad, and S. S. Ying, "Construction of Low Authority, Nearly Non-Minimal LQG Compensators for Reduced-Order Control Design," submitted to *IEEE Transactions on Automatic Control.* <Contained in Appendix I.>

2.10 E. G. Collins, Jr., S. S. Ying, W. M. Haddad, and S. Richter, "An Efficient, Numerically Robust Homotopy Algorithm for H_2 Model Reduction Using the Optimal Projection Equations," submitted to *International Journal of Control.* <Contained in Appendix J.>

2.11 E. G. Collins, Jr., W. M. Haddad, and S. S. Ying, "Reduced-Order Dynamic Compensation Using the Hyland and Bernstein Optimal Projection Equations," submitted to *Journal Guidance, Control, and Dynamics.* <Contained in Appendix K.>

2.12 W. M. Haddad, E. G. Collins, Jr., and R. Moser, "Computation of the Complex Structured Singular Value Using Fixed Structure Dynamic D-Scales," submitted to *System and Control Letters.* <Contained in Appendix G.>

2.13 W. M. Haddad, E. G. Collins, Jr., and R. Moser, "Structured Singular Value Controller Synthesis Using Constant D-Scales without D-K Iteration," submitted to *International Journal of Control.* <Contained in Appendix L.>

2.14 W. M. Haddad, E. G. Collins, Jr., D. C. Hyland, C.-S. Chellaboina, "New Frequency Domain Performance Bounds for Uncertain Structural Systems with Actuater and Sensor Dynamics," submitted to *Automatica.* <Contained in Appendix M.>

Conference Publications .

2.15 D. S. Bernstein, W. M. Haddad, D. C. Hyland, and F. Tyan, " A Maximum Entropy-Type Lyapunov Function for Robust Stability and Performance Analysis," *Proceedings of the American Control Conference*, Chicago, IL, pp. 355-356, June 1992.

2.16 E. G. Collins, Jr., L. D. Davis, and S. Richter, "A Homotopy Algorithm for Maximum Entropy

Design," *Proceedings of the American Control Conference*, Chicago, IL, pp. 1010-1014, June 1993.

2.17 W. M. Haddad and D. S. Bernstein, "The Multivariable Parabola Criterion for Robust Controller Synthesis: A Riccati Equation Approach," *Proceedings of the American Control Conference*, Chicago, IL, June 1992.

2.18 W. M. Haddad, E. G. Collins, Jr., and D. S. Bernstein, "Robust Stability Analysis Using the Small Gain, Circle, Positivity, and Popov Theorems: A Comparative Study," *Proceedings of the American Control Conference*, Chicago, IL, pp. 2425-2426, June 1992.

2.19 E. G. Collins, Jr., W. M. Haddad and L. D. Davis, "Riccati Equation Approaches for Robust Stability and Performance Analysis Using the Small Gain, Positivity, and Popov Theorems," *Proceedings of the American Control Conference*, San Francisco, CA, pp. 1079-1083, June 1993.

2.20 D. C. Hyland, E. G. Collins, Jr., W. M. Haddad, and V. S. Chellaboina, "Frequency Domain Performance Bounds for Uncertain Positive Real Plants Controlled By Strictly Positive Real Compensators," *Proceedings of the 1994 American Control Conference*, Baltimore, MD, pp. 2328-2332, June 1994.

2.21 J. P. How, E. G. Collins, Jr., and W. M. Haddad, "Optimal Popov Controller Analysis and Synthesis for Systems with Real Parameter Uncertainty," submitted to *the 1995 AIAA Navigations and Control Conference.*

2.22 E. G. Collins, Jr., W. M. Haddad, and S. S. Ying, "Generalized Fixed-Structure Optimality Conditions for H_2 Optimal Control," *Proceedings of the American Control Conference*, San Francisco, CA, pp. 2439-2443, June 1993.

2.23 E. G. Collins, Jr., W. M. Haddad, and S. S. Ying, "Construction of Low Authority, Nearly Non-Minimal LQG Compensators for Reduced-Order Control Design," *Proceedings of the 1994 American Control Conference*, Baltimore, MD, pp. 3411-3415, June 1994.

2.24 E. G. Collins, Jr., S. S. Ying, W. M. Haddad, and S. Richter, "An Efficient, Numerically Robust Homotopy Algorithm for H_2 Model Reduction Using the Optimal Projection Equations," *Proceedings of the 1994 IEEE Conference on Decision and Control*, to appear.

2.25 E. G. Collins, Jr., W. M. Haddad, and S. S. Ying, "Reduced-Order Dynamic Compensation Using the Hyland and Bernstein Optimal Projection Equations," to be submitted to *the 1995 American Control Conference.*

2.26 W. M Haddad, E. G. Collins, Jr., and R. Moser, "Fixed Structure Computation of the Structured Singular Value," *Proceedings of the American Control Conference*, San Francisco, CA, pp. 1010-1014, June 1993.

2.27 W. M. Haddad, E. G. Collins, Jr., and R. Moser, "Computation of the Complex Structured Singular Value Using Fixed Structure Dynamic D-Scales," *Proceedings of the 1994 IEEE Conference on Decision and Control*, to appear.

2.28 W. M. Haddad, E. G. Collins, Jr., and R. Moser, "Structured Singular Value Controller Synthesis Using Constant D-Scales without D-K Iteration," *Proceeding of the 1994 American Control Conference*, Baltimore, MD, pp. 2798-2802, June 1994.

2.29 W. M. Haddad, E. G. Collins, Jr., D. C. Hyland, and V. -S. Chellaboina, "New Frequency Domain Performance and Bounds for Uncertain Structural Systems with Actuater and Sensor Dynamics," submitted to *the 1995 American Control Conference*

3. Frequency Domain Performance Bounds for Uncertain Positive Real Plants Controlled by Strictly Positive Real Compensators [2.6, 2.14, 2.20, 2.29]

Many of the developments in robustness analysis have focused exclusively on the determination of stability. However, in practical engineering, performance issues are paramount, so it is also important to determine the type of performance degradation that occurs due to uncertainty in the system modeling. A common feature of a class of these results [3.1–3.4] is that they rely on majorant bounding techniques.

In [3.1–3.4] performance bounding is measured in basically three ways. References [3.1] and [3.2] measure performance in terms of second order statistics. In particular, bounds are obtained on the steady-state variances of selected system variables. In [3.3], performance is expressed in terms of the frequency response of selected system outputs. This result led to a new upper bound for the complex structured singular value [3.5]. Finally, [3.4] considers the transient response of certain system outputs, a performance measure which had not previously been treated in the robustness literature.

A common feature of these results and most other robustness results, with the possible exception of methods based on extensions of Popov analysis and parameter-dependent Lyapunov functions is that they do *not* predict unconditional stability for feedback systems consisting of a positive real plant controlled by a strictly positive real controller.

This research uses the logarithmic norm in the context of majorant analysis to develop tests for robust stability and performance that predict unconditional stability for the above case and also yield robust performance bounds. As in [3.3], this result considers the frequency domain behavior of a given system. The results are specialized to the case of static, decentralized, collocated rate feedback and dynamic, collocated rate feedback. An example of the results is shown in Figure 3.1 which shows the performance envelope predicted by the (new) positive real majorant analysis and the actual variations due to perturbations in the lowest natural frequency. For this case, completely norm-based majorant analysis [3.3] and complex structured singular value analysis [3.5] predicted instability. Figure 3.2 compares the positive real majorant bound (PRMB) with the complex block-structure majorant bound (CBSMB) from [3.3] and the complex structured singular value bound (CSSVB) derived from [3.5] for analysis of an Euler-Bernouilli beam with frequency uncertainty. Notice that over all frequencies PRMB < CBSMB < CSSVB.

These results have been extended to plants with sensor and actuater dynamics. Examples have shown similar nonconservatism to that described above.

References

3.1 D. C. Hyland and D. S. Bernstein, "The Majorant Lyapunov Equation: A Nonnegative Matrix Equation for Guaranteed Robust Stability and Performance of Large Scale Systems," *IEEE Transactions on Automatic Control*, Vol. AC-32, 1987, pp. 1005-1013.

3.2 E. G. Collins, Jr. and D. C. Hyland, "Improved Robust Performance Bounds in Covariance Majorant Analysis," *International Journal of Control*, Vol. 50, No. 2, 1989, pp. 495-509.

3.3 D. C. Hyland and E. G. Collins, Jr., "An M-Matrix and Majorant Approach to Robust Stability and Performance Analysis for Systems with Structured Uncertainty," *IEEE Transactions on Automatic Control*, Vol. 34, 1989, pp. 691-710.

3.4 D. C. Hyland and E. G. Collins, Jr., "Some Majorant Robustness Results for Discrete-time Systems," *Automatica*, Vol. 27, No. 1, 1991, pp. 167-172.

3.5 A. Packard and J. C. Doyle, "The Complex Structured Singular Value," *Automatica*, Vol. 29, 1993, pp. 71-109.
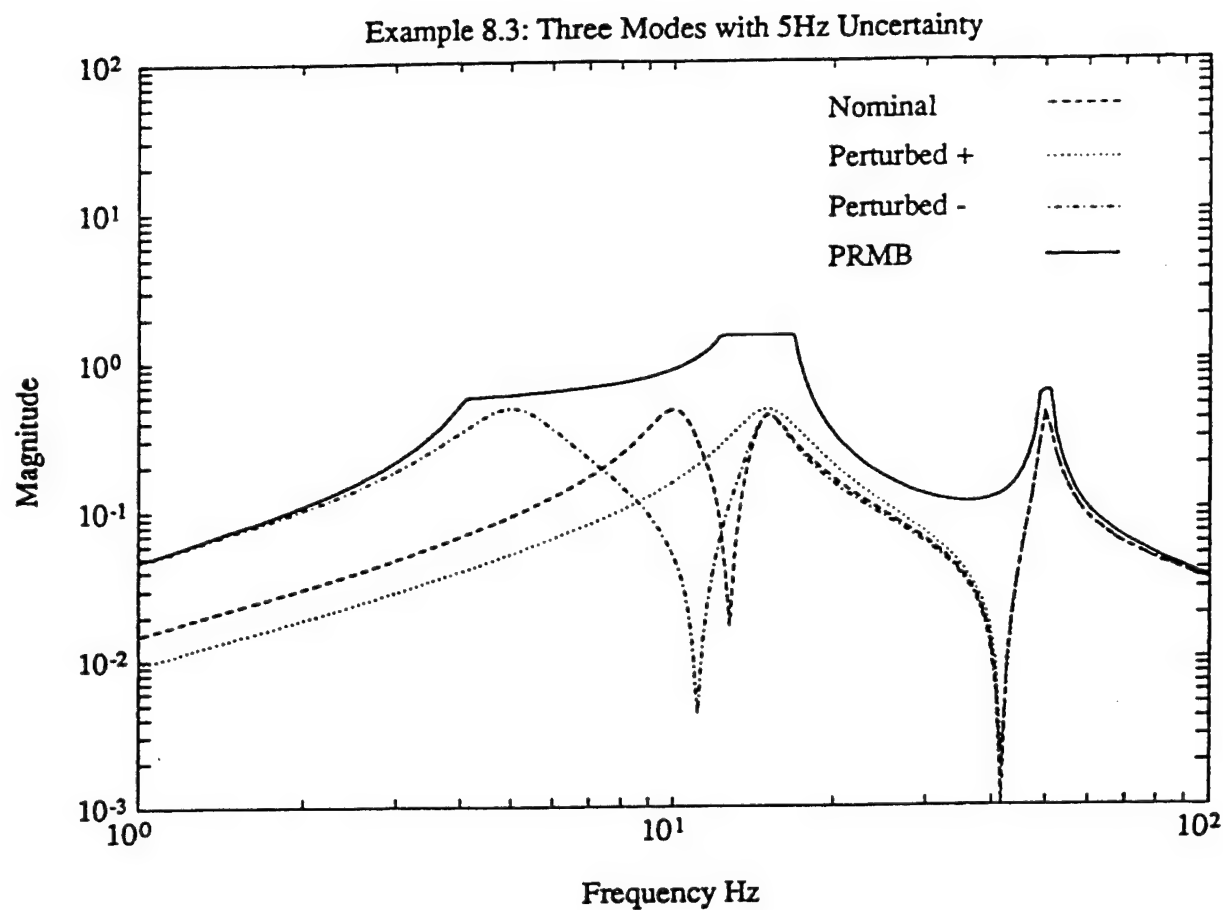
Figure 3.1. Performance Bound for Example 8.3 of [2.1]

(3 modes, lowest frequency uncertain)

Example 8.5: Euler-Bernoulli Beam
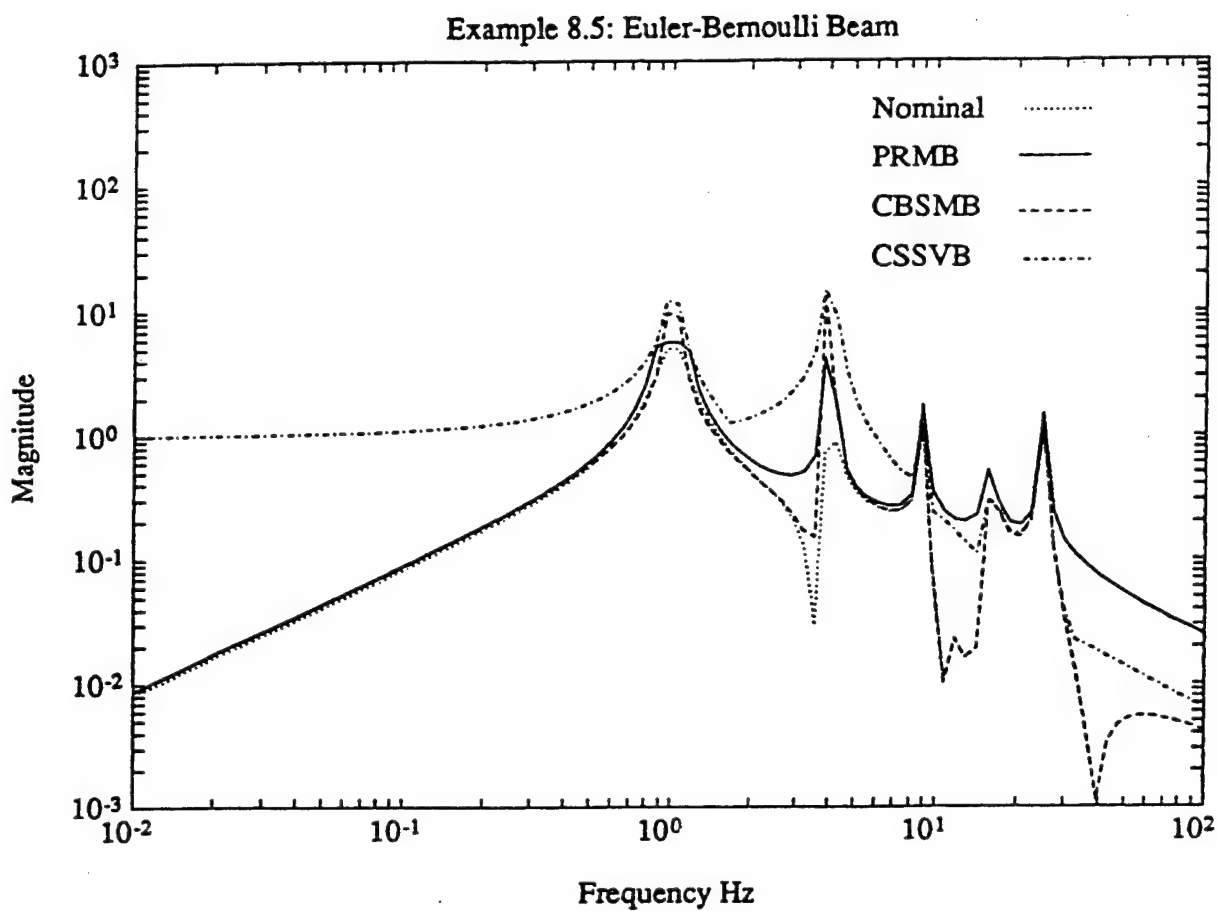
Figure 3.2. Comparison of PRMB, CBSMB, and CSSVB for Example 8.5 of [2.1]

4. Maximum Entropy-Type Lyapunov Functions for Robust Stability and Performance Analysis [2.1, 2.15]

The Maximum Entropy approach to robust control was developed to address the problem of modal uncertainty in flexible structures [4.1–4.4]. The rationale for this approach was based upon insights from the statistical analysis of lightly damped structures. Despite favorable comparisons to other approaches and experimental application, the basis and meaning of the approach remains mostly empirical. This research was initiated to make significant progress towards developing a rigorous foundation for Maximum Entropy design.

Besides statistical modal analysis techniques, a variety of formulations have been put forth for justifying the Maximum Entropy approach. To reproduce certain covariance phenomena of uncertain multimodal systems (decorrelation, incoherence, and equipartition) a multiplicative white noise model was invoked [4.1, 4.2]. The specific model chosen was interpreted in the sense of Stratonovich, thus entailing a critical correction term in the covariance equation due to the conversion from Stratonovich to Ito calculus. The Stratonovich model was itself based upon a limiting process in which the parameter entropy increased, thus suggesting the name "Maximum Entropy" control.

An alternative justification for the Maximum Entropy model was given in [4.5] where a covariance averaging approach was used to show that if the state covariance is averaged over uncertain modal frequencies possessing a Cauchy distribution, then the resulting averaged covariance satisfies the Maximum Entropy covariance model.

Although the various formulations of Maximum Entropy theory lend considerable insight into the nature of the approach, there remains a significant gap between this approach and more conventional techniques, such as H_∞ theory. The missing link, in our opinion, is the lack of a Lyapunov function that guarantees the robust stability of the closed-loop control system. In this regard it was long suspected that such a Lyapunov function would be unconventional, that is, unlike those arising in H_∞ theory. This view arose from the fact that the Maximum Entropy controllers were often robust to large perturbations in the damped natural frequencies, that is, the imaginary part of the eigenvalues. Such perturbations are highly structured, and thus are often treated conservatively by conventional small-gain-type bounds.

This research provided a Lyapunov-function basis for the Maximum Entropy covariance model for the case of modal frequency uncertainty. In fact, in this special case, two alternative Lyapunov

functions along with corresponding performance bounds were provided. Each Lyapunov function involves the sum of two matrices, the first being the solution to the Maximum Entropy equation, and the second being a constant auxiliary portion. The construction is similar to the parameter-dependent Lyapunov function technique developed in [4.6], except that in the present case the auxiliary portion is constant, that is, independent of the uncertainty.

While this research potentially provides a Lyapunov function foundation for the Maximum Entropy control approach, our results are currently limited to open-loop analysis. An illustration of the performance bounds predicted by the Maximum Entropy Lyapunov functions is given in Figure 4.1 which considers a one mode system with frequency uncertainty. Note that both Maximum Entropy performance bounds are much less conservative than the performance bound developed in [4.7].

References

4.1 D. C. Hyland and A. N. Madiwale, "A Stochastic Design Approach for Full-Order Compensation of Structural Systems with Uncertain Parameters," *Proceedings of the AIAA Guidance and Control Conference*, Albuquerque, NM, 1981, pp. 324–332.

4.2 D. C. Hyland, "Maximum Entropy Stochastic Approach to Controller Design for Uncertain Structural Systems," *Proceeding of the American Control Conference*, Arlington, VA, June 1982, pp. 680–688.

4.3 D. S. Bernstein and D. C. Hyland, "The Optimal Projection/Maximum Entropy Approach to Designing Low-order, Robust Controllers for Flexible Structures," *Proceedings of the IEEE Conference on Decision and Control*, Fort Lauderdale, FL, December 1985, pp. 745–752.

4.4 D. S. Bernstein and D. C. Hyland, "The Optimal Projection Approach to Robust, Fixed-Structure Control Design," in *Mechanics and Control of Space Structures*, J. L. Junkins, Ed., AIAA, 1990, pp. 287–293.

4.5 S. R. Hall, D. G. MacMartin and D. S. Bernstein, "Covariance Averaging in the Analysis of Uncertain Systems," *Proceedings of the American Control Conference*, Arlington, VA, June 1982, pp. 680–688.
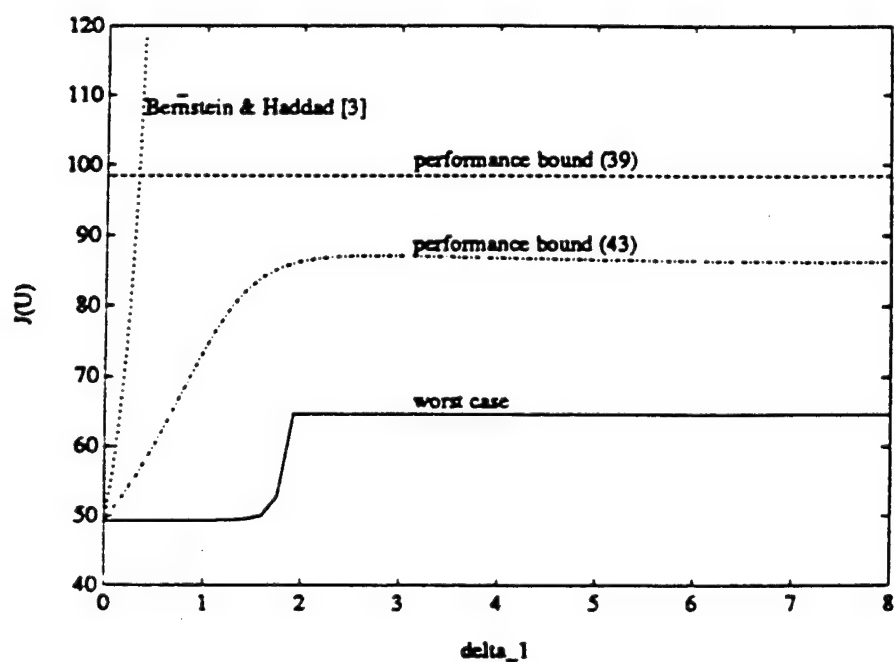
Figure 4.1. The performance bounds based on the Maximum Entropy Lyapunov functions (labeled (39) and (43)) were much less conservative than the performance bound based on the results of [4.7] (labeled Bernstein & Haddad)

4.6 W. M. Haddad and D. S. Bernstein, "Parameter-Dependent Lyapunov Functions, Constant Real Parameter Uncertainty, and the Popov Criterion in Robust Analysis and Synthesis," *Proceedings of the IEEE Conference on Decision and Control*, Brighton, U. K., December 1991, Part I, pp. 2274–2279, Part II, pp. 2632–2633.

4.7 D. S. Bernstein and W. M. Haddad, "Robust Stability and Performance Analysis for Linear Dynamic Systems," *IEEE Transactions on Automatic Control*, Vol. 34, 1989, pp. 751–758.

5.   A Homotopy Algorithm for Maximum Entropy Design [2.2, 2.16]

The linear-quadratic-guassian (LQG) compensator has been developed to facilitate the design of control laws for complex, multi-input multi-output (MIMO) systems such as flexible structures. However, it is well known that an LQG compensator can yield a closed-loop system with arbitrarily poor performance robustness properties. This deficiency has led to generalizations of LQG that allow the design of robust controllers. One such generalization of LQG is the Maximum Entropy control design approach discussed in the previous section. Although, as previously mentioned, the rigorous theoretical foundation for Maximum Entropy design is not yet complete, it has proven to be an effective tool in the design of robust control laws for ground-based flexible structure testbeds [5.1, 5.2] and for certain benchmark problems [5.3, 5.4].

The computation of full-order Maximum Entropy controllers requires the solution of a set of equations consisting of two Riccati equations coupled to two Lyapunov equations. If the uncertainty is assumed to be zero, these equations decouple and the Riccati equations become the standard LQG Riccati equations. A homotopy algorithm for solving these equations is described in [5.5]. This algorithm is based on first solving an LQG problem and gradually increasing the uncertainty level until the desired degree of robustness is achieved. Unfortunately, the algorithm of [5.5] relies on an iterative scheme that tends to have increasingly poor convergence properties as the uncertainty level is increased.

The contribution of this research is the development of a new homotopy algorithm for full-order Maximum Entropy design. The algorithm development utilizes the results of [5.6]. Unlike the previous approach, this algorithm has quadratic convergence rates along the homotopy curve. The algorithm has been implemented in MATLAB and is illustrated using a single-input, single-output control problem for the ACES testbed at NASA Marshall Space Flight Center in Huntsville, Alabama.

The Bode plots of the 17th-order open loop ACES plant are shown in Figure 5.1. The basic control objective is to attenuate the lower frequency modes of the structure (i.e., the modes less than 3 Hz). Each of the flexible modes is considered uncertain. The magnitude of the uncertainties is determined by a scalar parameter (β).
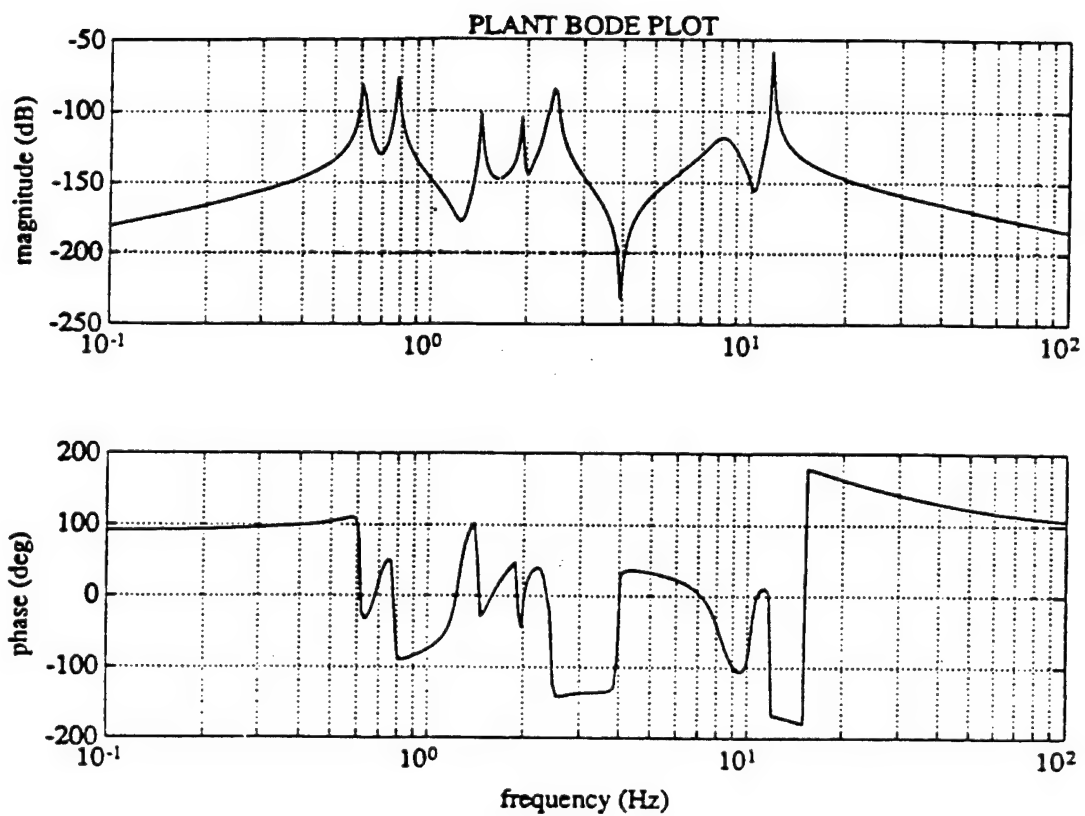
Figure 5.1. The SISO ACES transfer function consisted of seven distinct modes
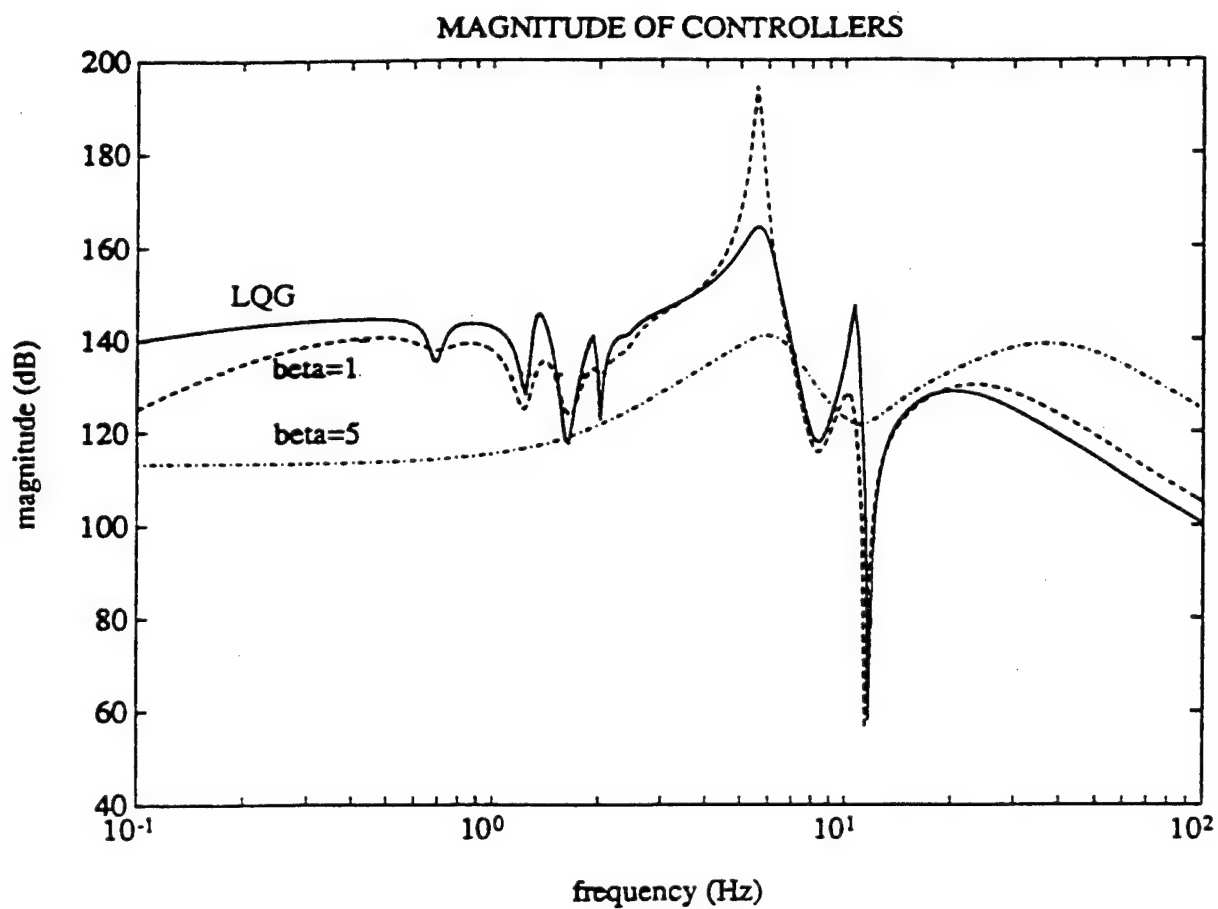
Figure 5.2. The frequency response of the Maximum Entropy Controllers became increasingly smooth as the uncertainty level (proportional to β) was increased
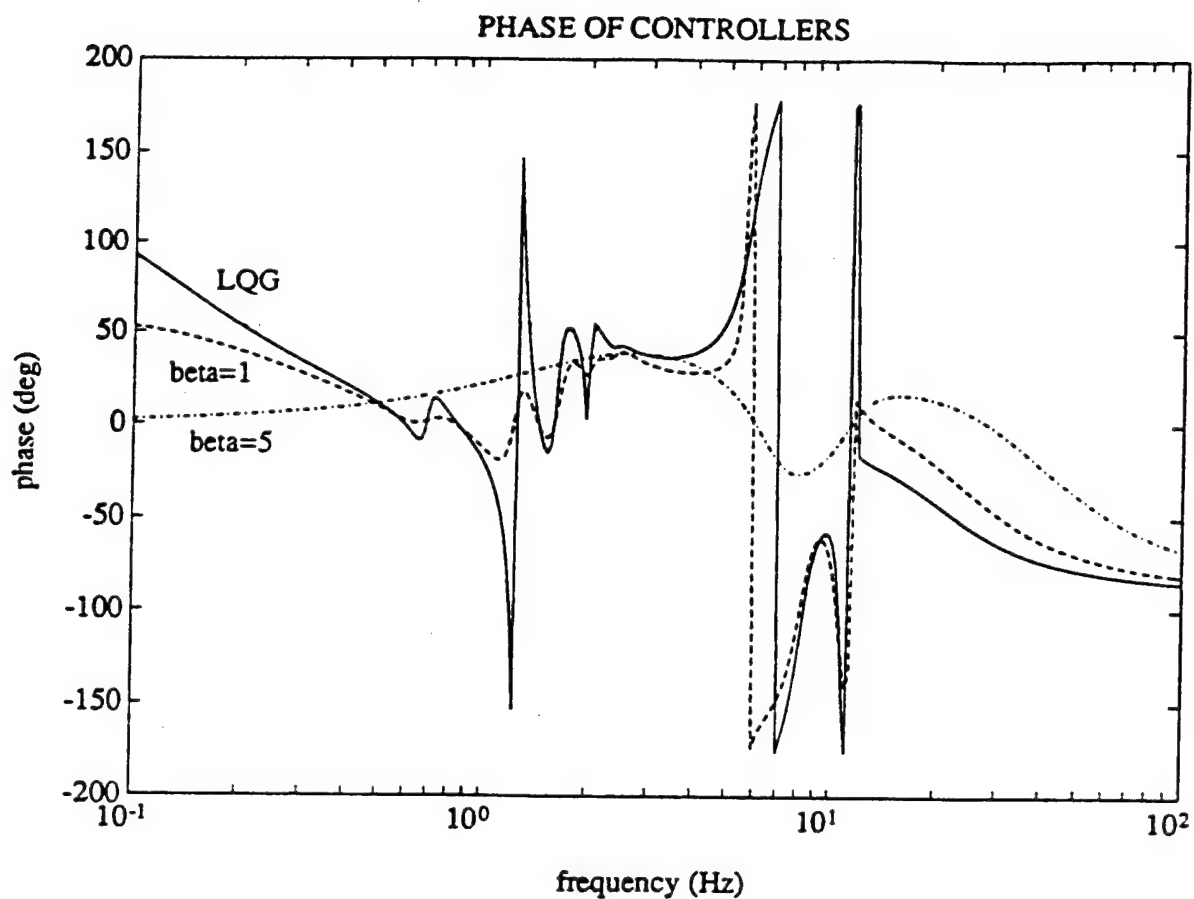
Figure 5.3. The phase frequency response of the Maximum Entropy controllers shows that the phase becomes positive real over a large frequency range as the uncertainty level (proportional to β) was increased

For this example, the MATLAB implementation of the Maximum Entropy Homotopy algorithm was run on a 486, 33 MHz PC. Table 5.1 shows some of the runtime statistics of the program. The highest uncertainty design, corresponding to $\beta=5$ was obtained in approximately one hour. Notice that the number of flops and the run time are essentially linear with respect to the log of the scale factor β. This general trend has also been observed in other design examples.

Table **5.1.** Run-Time Statistics of the Maximum Entropy Homotopy Algorithm

Initial beta	Final beta	Megaflops	RealTime (sec.)	Predictions & Corrections
0	.01	1246.25	1027.27	43
0.1	.1	1061.41	884.80	36
.1	1	1061.49	889.84	36
1	5	1083.25	995.87	41

Figures 5.2 and 5.3 compare respectively the magnitude and phase of the initial LQG controller and the Maximum Entropy controllers corresponding to $\beta = 1$ and $\beta = 5$. Notice that the $\beta = 5$ controller has a very smooth frequency response and is positive real over a very large frequency band. The smoothness of this controller indicates that its effective order is much less than 17. Using balanced controller reduction, a 4th order compensator was obtained that was nearly identical to the 17th order compensator.

References

5.1 E. G. Collins, Jr., D. J. Phillips, and D. C. Hyland, "Robust Decentralized Control Laws for the ACES Structure," *Control Systems Magazine*, Vol. 11, April 1991, pp. 62–70.

5.2 E. G. Collins, Jr., J. A. King, D. J. Phillips, and D. C. Hyland, "High Performance, Accelerometer-Based Control of the Mini- MAST Structure," *AIAA J. Guidance Control and Dynamics*, Vol. 15, July 1992, pp. 885–892.

5.3 M-F Cheung and S. Yurkovich, "On the Robustness of MEOP Design Versus Asymptotic LQG Synthesis," *IEEE Transactions on Automatic Control*, Vol. 33, November 1988, pp. 1061–1065.

5.4 E. G. Collins, Jr., J. A. King, and D. S. Bernstein, "Application of Maximum Entropy/Optimal Projection Design Synthesis to a Benchmark Problem," *Journal of Guidance, Control and Dynamics*, Vol. 15, July 1992, pp. 885–892.

5.5 E. G. Collins, Jr. and S. Richter, "A Homotopy Algorithm for Synthesizing Robust Controllers for Flexible Structures Via the Maximum Entropy Design Equations," *Third Air Force/NASA Symposium on Recent Advances in Multidisciplinary Analysis and Optimization*, San Diego,

CA, May 1990, pp. 1449–1454.

5.6 S. Richter, L. D. Davis, and E. G. Collins, Jr., "Efficient Computation of the Solutions to Modified Lyapunov Equations," *SIAM Journal of Matrix Analysis and Applications, January 1993.*

6. The Multivariable Parabola Criterion for Robust Control Design and Analysis [2.13, 2.17]

One of the most basic issues in system theory is the stability of feedback interconnections. Four of the most fundamental results concerning stability of feedback systems are the small gain, positivity, circle and Popov theorems. In a recent paper [6.1], each result was specialized to the problem of robust stability involving linear uncertainty, and a Lyapunov function framework was established providing connections between these classical results and robust stability via state space methods. As shown in [6.1], the main difference between the small gain, positivity, and circle theorems versus the Popov theorem is that the former results guarantee robustness with respect to arbitrarily time-varying uncertainty while the latter does not. This is not surprising since the Lyapunov function foundation of the small gain, positivity, and circle theorems is based upon conventional or "fixed" quadratic Lyapunov functions which guarantee stability with respect to arbitrarily time-varying perturbations. Since time-varying parameter variations can destabilize a system even when the parameter variations are confined to a region in which constant variations are nondestabilizing, a feedback controller designed for time-varying parameter variations may unnecessarily sacrifice performance when the uncertain real parameters are constant.

Whereas the small gain, positivity and circle results are based upon fixed quadratic Lyapunov functions, the Popov result is based upon a quadratic Lyapunov function that is a *function* of the parametric uncertainty. Thus, in effect, the Popov result guarantees stability by means of a *family* of Lyapunov functions. For robust stability, this situation corresponds to the construction of a parameter-dependent quadratic Lyapunov function [6.2]. A key aspect of this approach is the fact that it does *not* apply to arbitrarily time-varying uncertainties, which renders it less conservative than fixed quadratic Lyapunov functions (such as the small gain, positivity, and circle results) in the presence of constant real parameter uncertainty. An immediate application of the parameter-dependent Lyapunov function framework of [6.2] is the reinterpretation and generalization of the classical Popov criterion as a parameter-dependent Lyapunov function for constant linear parametric uncertainty.

From a theoretical perspective, an important contribution of this research is the unification of the circle and Popov criteria via a parameter-dependent Lyapunov function framework that yields both results as special cases. The unification of the circle and Popov criteria per se is not new to this research. Indeed, a parablola test which accomplishes this goal was originally developed in [6.3] and further studied in [6.4]. However these results are confined to SISO systems and rely on

graphical techniques. This research thus accomplished four specific goals:

1. It provided a general framework for the parabola test in terms of parameter-dependent Lyapunov functions in the spirit of [6.2];

2. It developed a state space characterization of the parabola test via Riccati equations;

3. It developed a *multivariable* extension of the parabola test for parametric uncertainty; and

4. It used these results to develop equations for robust controller synthesis.

One of the limitations of Popov theory is that it restricts the uncertainty to vary in only one direction (that is, positive or negative). The parabola test, however, allows the uncertainty to vary in both directions and hence can potentially lead to analysis and design tools that are more easily applied than those resulting from Popov theory. Hence, this research could result in robustness analysis tools that are more useful to the practicing controls engineer.

References

6.1 W. M. Haddad and D. S. Bernstein, "Explicit Construction of Quadratic Lyapunov Functions for the Small Gain, Positivity, Circle, and Popov Theorems and Their Application to Robust Stability," in *Control of Uncertain Dynamic Systems*, S. P. Bhattacharyya and L. H. Keel, Eds., CRC Press, pp. 149-173, 1991.

6.2 W. M. Haddad and D. S. Bernstein, "Parameter-Dependent Lyapunov Functions, Constant Real Parameter Uncertainty, and the Popov Criterion in Robust Analysis and Synthesis," *Proceedings of the IEEE Conference on Decision and Control*, Brighton, U. K., December 1991, Part I, pp. 2274-2279, Part II, pp. 2632-2633.

6.3 K. S. Narendra and J. H. Taylor, *Frequency Domain Criteria for Absolute Stability*, Academic Press, New York, 1973.

6.4 A. R. Bergen and M. A. Sapiro, "The Parabola Test for Absolute Stability," *IEEE Transactions on Automatic Control*, Vol. AC-12, pp. 312-314, 1967.

7. Application of Popov Robustness Tests to a Benchmark Problem [2.4, 2.5, 2.18, 2.19]

Over the past several years, significant attention has been devoted to the use of small gain(or H_∞) tests for robustness analysis. However, it is well known that these tests can be very conservative since in the frequency domain the small gain test characterizes uncertainty with bounded gain but arbitrary phase while in the time domain the small gain test characterizes uncertainty with arbitrary time variation. This conservatism has led to the search for more accurate robustness tests. In particular, researchers have sought tests that allow frequency domain uncertainty characterization to include phase bounding or time domain uncertainty characterization to include restrictions on the allowable time variations.

As discussed in the previous section, the small gain, circle and positivity tests are based upon conventional or "fixed" quadratic Lyapunov functions which guarantee stability with respect to arbitrarily time-varying perturbations. In contrast, the Popov test, based on a parameter dependent Lyapunov function, restricts the allowable time variation of the perturbation.

In this research we used a benchmark problem to compare the Popov test with the small gain and positivity tests. Each of the stability tests have graphical interpretations for the case of one block, scalar uncertainty. These graphical tests were applied. However, the state space tests that are based on Riccati equations are emphasized since they extend to more general forms of uncertainty and also allow the development of robust H_2 performance bounds. Homotopy algorithms were developed for the special case of one-block, scalar uncertainty. The algorithm for Popov analysis additionally required that a certain product $(\tilde{C}_0 \tilde{B}_0)$ related to the uncertainty characterization equal zero. This condition does hold for the benchmark problem under consideration. The robustness tests were applied to analyze a feedback system for the benchmark system in which the controller was designed using the Maximum Entropy approach.

The open-loop benchmark system is the two-mass/spring system shown in Figure 7.1. The stiffness k is uncertain. A control force acts on body 1, and the position of body 2 is measured, resulting in a noncolocated control problem. Here, we consider Controller #1 of [7.1] which was designed for Problem #1 of a benchmark problem [7.2] using the Maximum Entropy robust control design technique. The controller was designed so that the closed-loop system is robust with respect
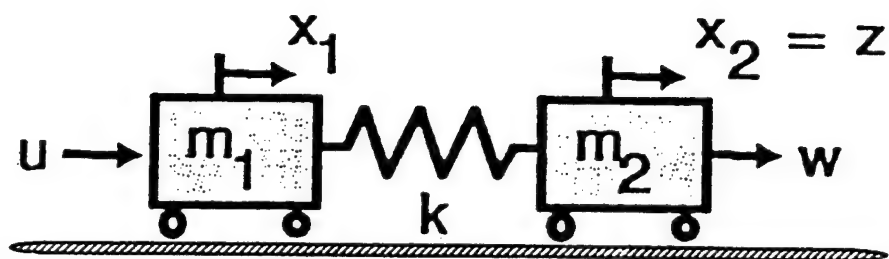
Figure 7.1. The benchmark system for robust control design and analysis is a two-mass/spring system

to perturbations in the nominal value of the stiffness k (i.e., $k = k_{nom}$). The exact stiffness stability region over which the system will remain stable was computed by a simple search and is given by

$$0.4458 < k < 2.0661.$$

Next, using a graphical approach and the state-space Riccati equation approach (implemented via homotopy algorithms), we apply small gain analysis, positivity analysis, and Popov analysis to determine the stiffness stability regions predicted by each of these tests. Each of these tests is related to the previous test and is guaranteed to be less conservative.

When the homotopy algorithms corresponding to the state space tests for small gain, positivity, and Popov analysis were applied to the benchmark problem, the performance curves shown in Figure 7.2 resulted. As expected, Popov analysis yielded less conservative results than the positivity and small gain tests. The robust stability bounds $\overline{\Delta k}$ (positive) and $\underline{\Delta k}$ (negative) obtained from the state space tests were identical to those obtained from the frequency domain tests. In fact, the stability region predicted by the Popov test was *identical* to the true stability region!

References

7.1 E. G. Collins, Jr., J. A. King, and D. S. Bernstein, "Application of Maximum Entropy/Optimal Projection Design Synthesis to a Benchmark Problem," *Journal of Guidance, Control and Dynamics*, Vol. 15, September 1992, pp. 1094–1102.

7.2 B. Wie and D. S. Bernstein, "Benchmark Problems for Robust Control Design," *Journal of Guidance, Control and Dynamics*, Vol. 15, September 1992, pp. 1057–1059.
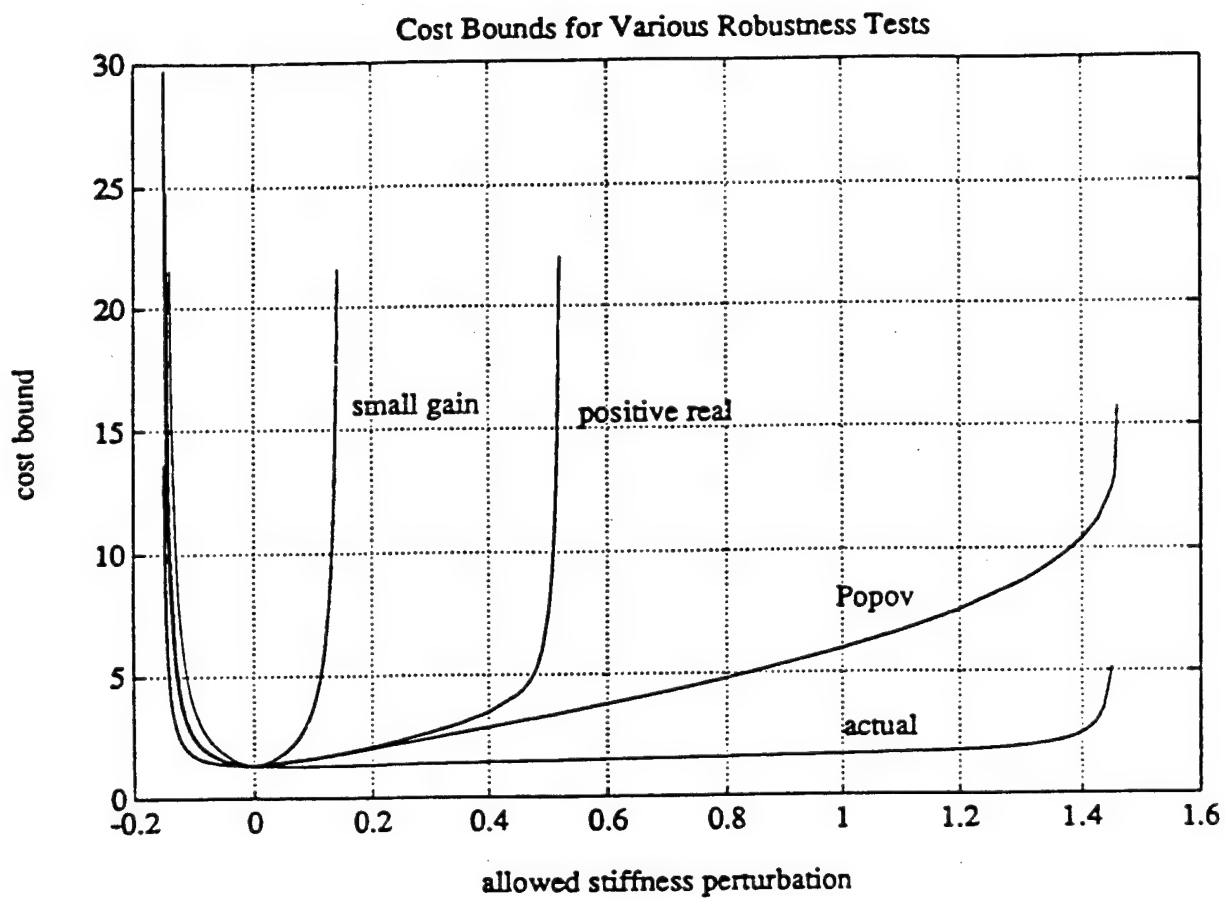
Figure 7.2. The performance bounds predicted by Popov analysis were significantly less conservative than those predicted by the small gain and positivity tests

8. A Numerical Algorithm for Optimal Popov Controller Analysis and Applications to a Structural Testbed [2.7, 2.21]

One of the most important aspects of the control design and evaluation process is the analysis of feedback systems for robust stability and performance. Significant attention has been devoted over the past several years to the use of bounded gain and other norm-based methods for these analysis tests. Unfortunately, due to their dependence on norms, these tests exclude the phase information on the system uncertainties and can be very conservative for systems with constant real parameter errors. A technique to reduce the conservatism inherent in fixed quadratic Lyapunov functions has recently been introduced (see, e.g., [2.3, 2.17]). The approach considers Lyapunov functions that explicitly contain the uncertain parameters, and thus restrict the allowable time-variation of the uncertainties.

The purpose of this research is to combine several recent advances on Popov controller analysis and synthesis. Refs. [2.5, 2.19] have recently demonstrated that the state space Popov analysis criterion is much less conservative than similar positive real and small gain (H_∞) criteria. In this research, we extend the earlier work by considering systems with multiple uncertainties that have both upper and lower sector bounds. The stability criterion is developed using a more general stability multiplier

$$W(s) = H + Ns, \qquad H > 0, \quad N \geq 0.$$

The algorithm of [2.5, 2.19] was developed for $H = I$. The new algorithm also considers the case $C_0 B_0 \neq 0$. The simplifying assumptions in [2.5, 2.19] that $C_0 B_0 = 0$ is only valid for a very restricted set of parameter uncertainties.

The optimal Popov analysis algorithm is demonstrated using several robust control designs that were developed for the Middeck Active Control Experiment (MACE) (see Figure 8.1) located at the Massachusetts Institute of Technology. Figure 8.2 shows the curves of robust (H_2) performance vs. guaranteed robust stability for an LQG controller, a Maximum Entropy (ME) controller and two multiple model (MM) controllers. Each of the controllers had at least one unstable eigenvalue except the "stable MM." Figure 8.3 shows the improvement in the stable MM design when it was refined using Popov controller synthesis (PCS).
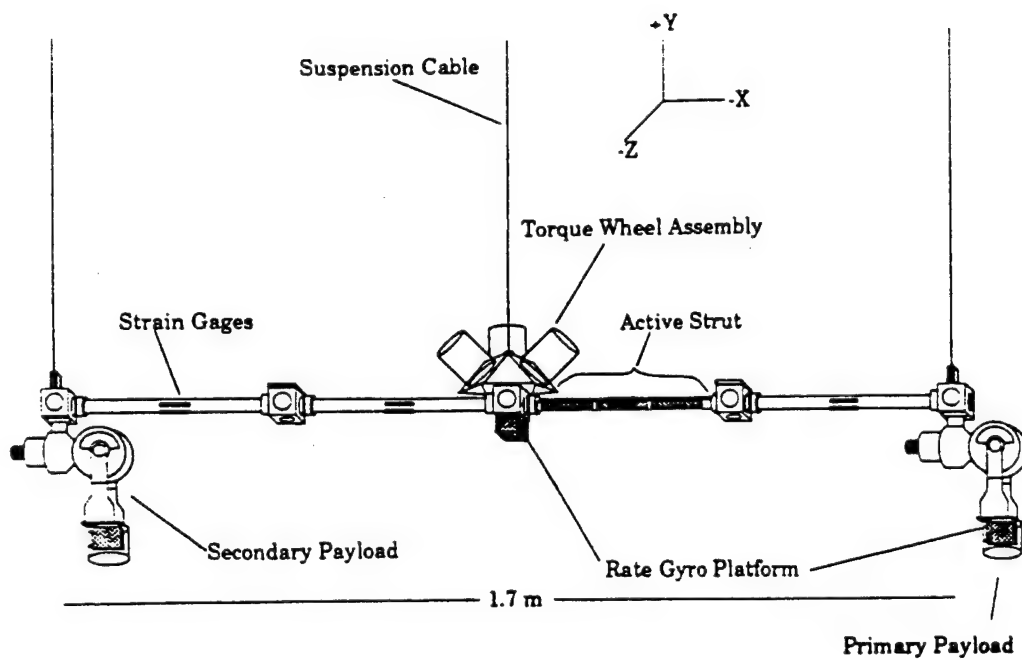
Figure 8.1. Middeck Active Control Experiment (MACE) Test Article
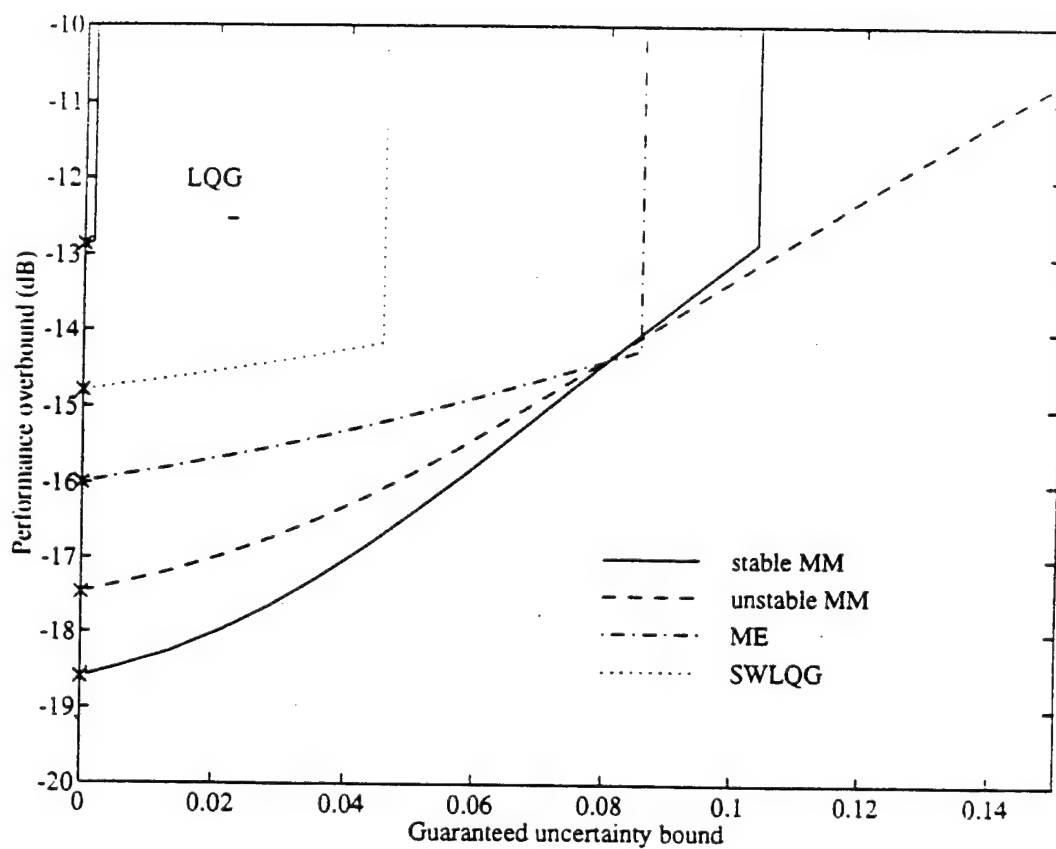
Figure 8.2. Robust Stability and Performance Analysis Using Several Controllers for MACE

(Symbols × indicate nominal H_2 performance for each design)
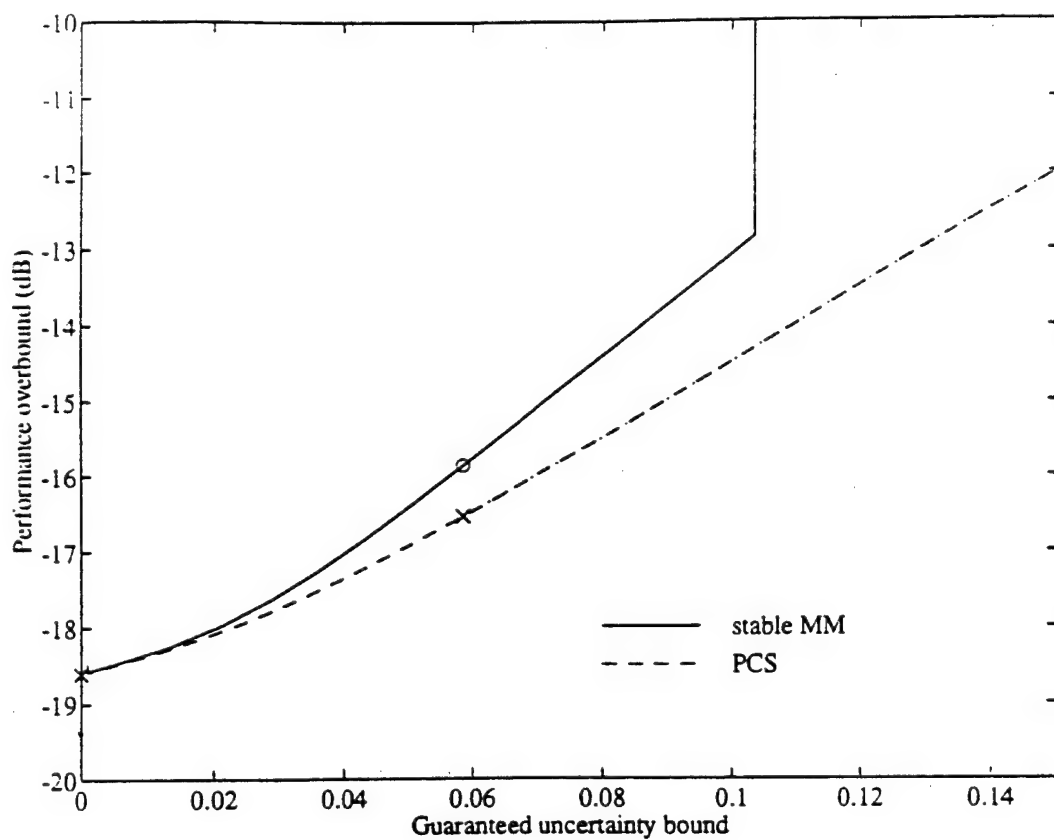
Figure 8.3. Robustness Improvements Achieved by Popov Controller Synthesis

9. Generalized Fixed-Structure Optimality Conditions for H_2 Optimal Control [2.8, 2.22]

One of the foundational results in modern control theory is the development of a characterization of the globally optimal H_2 controller via algebraic Riccati equations. This result has traditionally been derived via the Calculus of Variations or the Maximum Principle in conjunction with the Separation Principle. Unfortunately, the optimal H_2 or LQG (Linear-Quadratic-Gaussian) controller has dimension equal to that of the plant (although it may have minimal dimension which is less than that of the plant). This has motivated the search for optimal reduced-order controllers (i.e., controllers that have dimension less than that of the plant).

Because the Calculus of Variations and the Maximum Principle characterize globally optimal solutions, these traditional methods for deriving the LQG result do not extend to the development of characterizations of optimal reduced-order controllers. Hence, researchers have developed the optimization methods that allow the dimension and structure of the controller to be constrained *a priori*. These methods are usually based on Lagrange multiplier theory and will be called here "fixed-structure approaches." The "optimal projection" characterization of the necessary conditions for optimal reduced-order control was derived using a fixed-structure approach and yields the standard LQG regulator and observer Riccati equations when the dimension of the controller is specified to be equal to the dimension of the plant. However, the original optimal projection results and numerous extensions were derived by *a priori* assuming that the controller is minimal. This is a limiting assumption since it is known that even an LQG controller is not always minimal.

This research develops optimality conditions that are derived without assuming the minimality of the compensator. The results are specialized to the case in which the compensator is constrained to have the dimension of the plant. It is shown that even when compensator minimality is not assumed, fixed-structure theory is able to derive the LQG Riccati equations. It is also shown that there exist sets of coupled Riccati and Lyapunov equations that are identical in form to the optimal projection equations for reduced-order control but actually characterize extremals to the full-order compensation problem. This leads to a new interpretation of an optimal projection controller. In particular, an optimal projection controller is a projection, described by a projection matrix μ, of a "central" extremal to the H_2 optimal full-order compensation problem.

These latter results are used to discuss suboptimal projection methods that are able to produce minimal order realizations of nonminimal LQG compensators. For this special case, the similarity

transformations relating the projection matrix ν used by these suboptimal methods to the projection matrix μ and the optimal projection matrix τ from the standard optimal projection theory are explicitly defined. If the observability and controllability grammians of the nonminimal LQG compensator satisfy certain rank conditions, the three projection matrices are proved to be identical (i.e., $\tau = \mu = \nu$).

10. Construction of Low Authority, Nearly Non- Minimal LQG Compensators for Reduced-Order Control Design [2.9, 2.23]

The development of linear-quadratic-guassian (LQG) theory was a major breakthrough in modern control theory since it provides a systematic way to synthesize high performance controllers for nominal models of complex, multi-input multi-output systems. However, as discussed above, one of the well known deficiencies of an LQG compensator is that its minimal dimension is usually equal to the dimension of the design plant. This has led to the development of techniques to directly synthesize optimal, reduced-order controllers and techniques to synthesize reduced-order approximations of the optimal full-order compensator (i.e., controller reduction methods).

The controller reduction methods almost always yield suboptimal (and sometimes destabilizing) reduced-order control laws since an optimal reduced-order controller is not usually a direct function of the parameters used to compute or describe the optimal full-order controller. Nevertheless, these methods are computationally inexpensive and sometimes do yield high performing and even nearly optimal control laws. An observation that holds true about most of these methods is that they tend to work best at low control authority. However, to date no rigorous explanation has been presented to explain this phenomenon.

One of the purposes of this paper is to provide a partial explanation as to why the suboptimal projection methods tend to work at low control authority. The discussion here focuses on stable systems. It is shown that if the state weighing matrix R_1 or disturbance intensity (or covariance for discrete systems) V_1 has a specific structure in a basis in which the A matrix is upper or lower block triangular, respectively, then, as illustrated by Figure 8.1, at low control authority the corresponding LQG compensator is nearly nonminimal and can hence be easily reduced to a nearly optimal reduced-order controller. The conditions presented for R_1 and V_1 often are satisfied or nearly satisfied in practice. Hence, for stable systems the results proved in this research do offer one explanation of why suboptimal controller reduction methods often provide nearly optimal control laws at low authority. If these conditions are *not* satisfied, then, as illustrated by Figure 10.2, at low control authority the LQG compensator is not necessarily nearly nonminimal. The basic results can be used as guidelines for choosing R_1 and V_1 such that suboptimal controller reduction methods yield "good" reduced-order controllers.
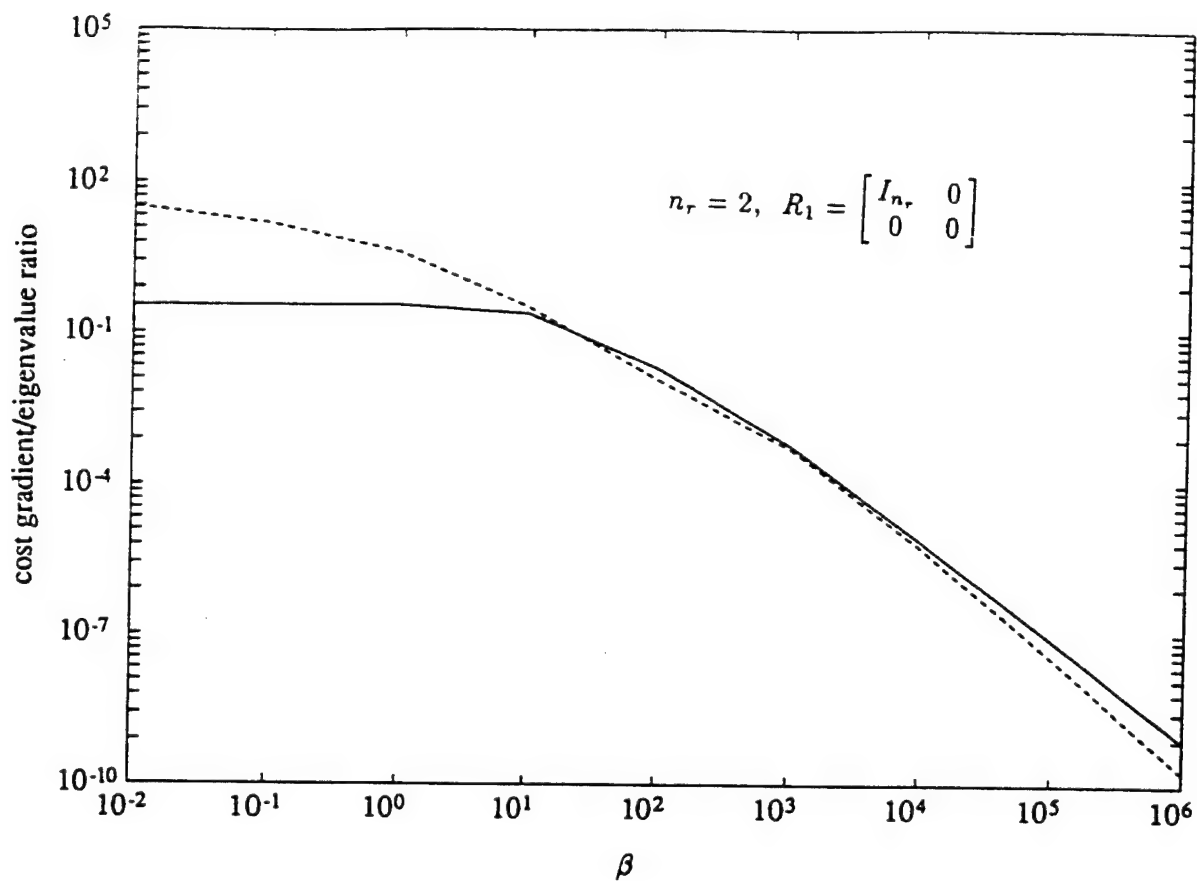
Figure 10.1. Non-minimality Indicator of the LQG Controller (solid line) and the Norm of the Cost Gradient of a 2nd-order Balanced Controller (dashed line) for a "<u>Good Structured</u>" R_1
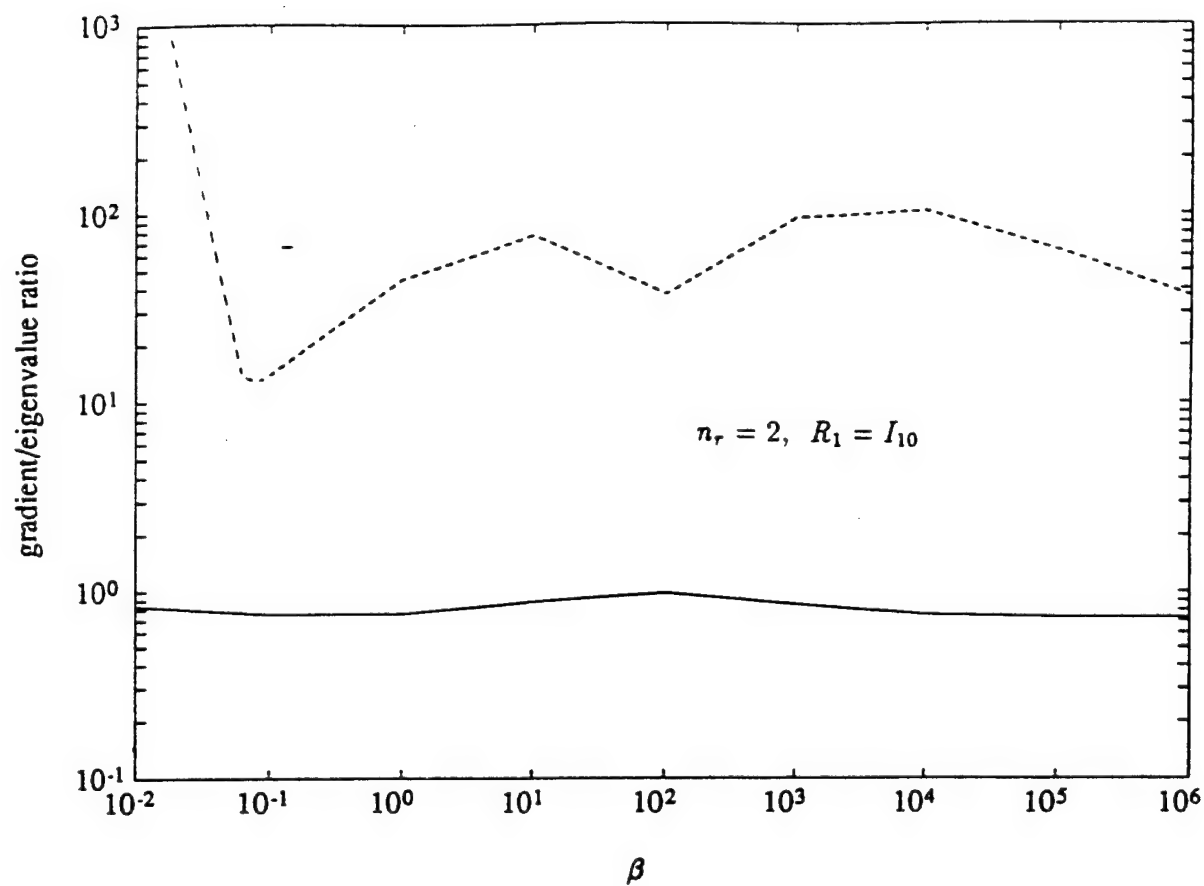
Figure 10.2. Non-minimality Indicator of the LQG Controller (solid line) and the Norm of the Cost Gradient of a 2nd-order Balanced Controller (dashed line) for a "<u>Bad Structured</u>" R_1

Suboptimal controller reduction methods can be used to initialize algorithms for synthesizing reduced-order controllers. Of particular interest are continuation and homotopy algorithms since they are based on allowing the plant and weights defining an optimization problem to vary as a function of the homotopy parameter $\lambda \in [0, 1]$. These homotopy algorithms rely on choosing the initial plant and weights so that the corresponding LQG compensator is easily reduced to a nearly optimal reduced-order compensator of the desired dimensions. Hence, the results developed in this research provide some rigorous guidelines for initializing these algorithms. Note that the restriction to stable systems is not necessarily limiting since the freedom involved in defining a continuation or homotopy map allows this assumption to be satisfied. However, future work will focus on theory that directly applies to unstable systems.

11. Continuation Algorithms for H_2 Optimal Reduced-Order Modeling and Control Using the Optimal Projection Equations [2.10, 2.11, 2.24, 2.25]

Most algorithms to date for H_2 optimal reduced-order modeling and control are descent algorithms, such that at each iteration they are guaranteed to decrease the cost. An exception has been the continuation and homotopy algorithms of [11.1–11.6]. These algorithms are not descent methods and since the shortest path from a given initial condition to an optimal solution is not necessarily a descent path, these algorithms have the potential to be more efficient than the descent methods. In addition, when a physical continuation or homotopy path is used, the reduced-order model or controller at each point along the homotopy path is guaranteed to be a meaningful model or controller for the physical system. Under mild conditions, the homotopy paths of the algorithms developed in [11.5, 11.6] are guaranteed to exist.

A common feature of the continuation and homotopy algorithms of [11.4–11.6] is that they are based directly on the gradient expressions. In these schemes, the parameter vector p represents the reduced-order model or controller. In order to keep the dimensionality of p relatively small and to avoid high order singularities along the homotopy path, minimal-order parameterizations of the reduced-order model or controller were considered. However, since the assumed parameterization may fail to exist or lead to ill-conditioning related to the insistence on using the minimal number of parameters, these resulting algorithms sometimes fail or have very poor convergence properties. On alternative approach proposed in [11.4–11.6] is to develop an algorithm that utilizes several minimal parameter homotopies and is capable of switching to an alternative parameterization if ill-conditioning is encountered with the current parameterization. A second approach is to develop algorithms directly based on the optimal projection equations.

Continuation and homotopy algorithms based on the optimal projection equations are given in [11.1–11.3]. However, the homotopy algorithms of [11.2, 11.3] suffer from the curse of large dimensionality. The continuation algorithm of [11.1] used a very crude path following scheme in which the coupled Riccati and Lyapunov equations comprising the optimal projection equations for reduced-order controller design were *not* updated simultaneously. This caused the algorithm to exhibit poor convergence properties, especially as the control authority was increased.

This research uses the optimal projection equations to develop new continuation algorithms for H_2 optimal, reduced-order modeling and control. These algorithms avoid the large dimensionality of [11.2, 11.3] by using the results of [11.7] to efficiently solve sets of linearly coupled Lyapunov

equations whose solutions describe either the tangent vectors or Newton corrections. The poor convergence properties of [11.1] are avoided by *simultaneously* updating each of the optimal projection equations. The new continuation algorithm for H_2 optimal reduced-order controller design produced the optimal curves for the benchmark "four disk problem" which are shown in Figure 11.1.

Note that the design model was 8th order. .

References

11.1 S. Richter and E. G. Collins, Jr., "A Homotopy Algorithm for Reduced-Order Controller Design Using the Optimal Projection Equations," *Proceedings of the IEEE Conference on Decision and Control*, Tampa, FL, pp. 506-511, December 1989.

11.2 D. Zigic, L. T. Watson, E. G. Collins, Jr., and D. S. Bernstein, "Homotopy Methods for Solving the Optimal Projection Equations for the H_2 Reduced Order Model Problem," *International Journal of Control*, Vol. 56, pp. 173-191, 1992.

11.3 D. Zigic, L. T. Watson, E. G. Collins, Jr., and D. S. Bernstein, "Homotopy Approaches to the H_2 Reduced Order Model Problem," *Journal of Mathematical Systems, Estimation, and Control*, to appear.

11.4 E. G. Collins, Jr., L. D. Davis, and S. Richter, "Design of Reduced-Order H_2 Optimal Controllers Using a Homotopy Algorithm," *International Journal of Control*, 1993, to appear.

11.5 Y. Ge, E. G. Collins, Jr., L. T. Watson, and L. D. Davis, "An Input Normal Form Homotopy for the L^2 Optimal Model Order Reduction Problem," *IEEE Transactions on Automatic Control*, to appear.

11.6 Y. Ge, E. G. Collins, Jr., L. T. Watson, and L. D. Davis, "A Comparison of Homotopies for Alternative Formulations of the L^2 Optimal Model Order Reduction Problem," submitted to *International Journal of Control*.

11.7 S. Richter, L. D. Davis, and E. G. Collins, Jr., "Efficient Computation of the Solutions to Modified Lyapunov Equations," *SIAM Journal of Matrix Analysis and Applications*, pp. 420-431, 1993.
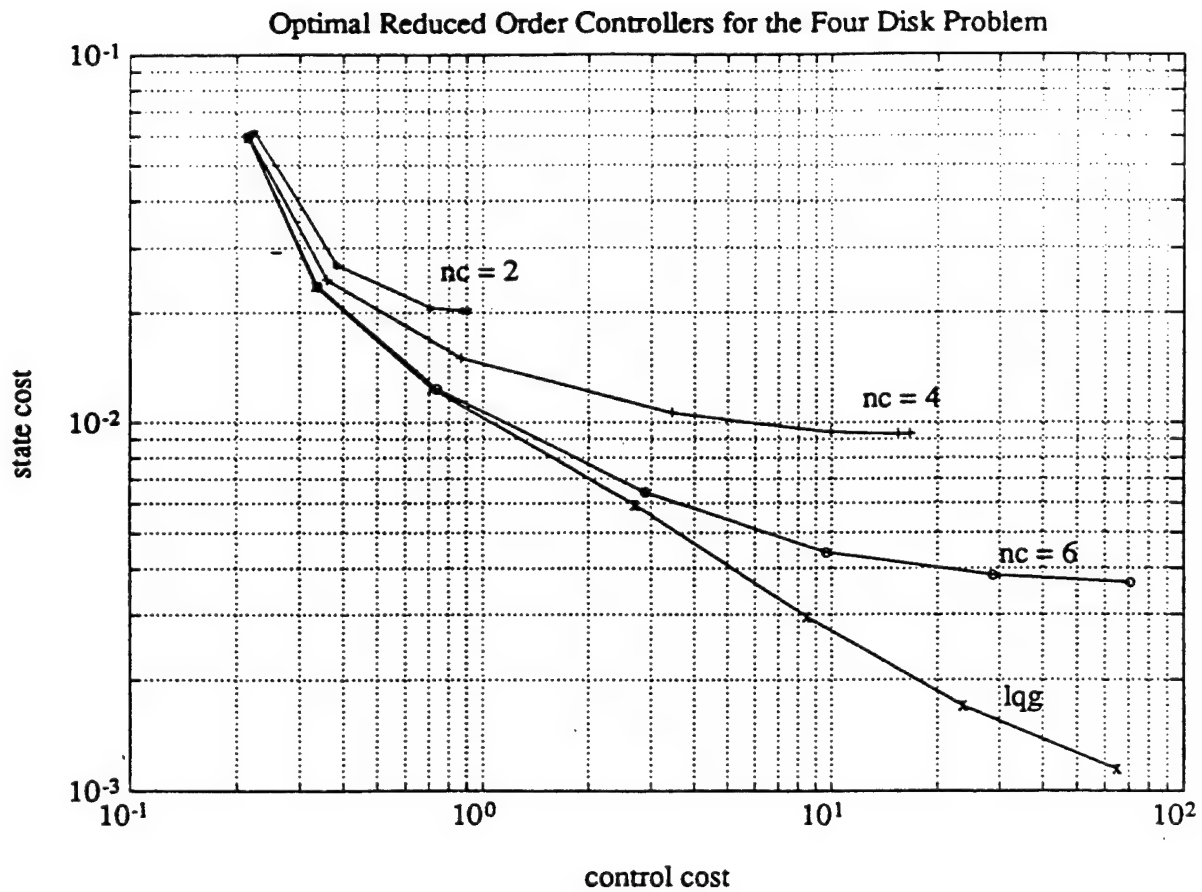
Figure 11.1. Comparison of the Performance Curves for Various Order Controllers for an 8th Order Four-Disk Plant

12. Analysis and Synthesis with the Complex Structured Singular Value Using Fixed Structure D-Scales [2.12, 2.13, 2.26–2.28]

A fundamental problem in control engineering is the design of feedback controllers that are insensitive to errors in the control design model. The characterization of the uncertainty occurs somewhere between two extremes, parametric and nonparametric uncertainty. Parametric uncertainty here describes errors that can be translated into errors in the elements of some time-invariant, state space representation of the design model. An example of this type of uncertainty would be errors in the mass or stiffness parameters of a finite element model. On the other hand, nonparametric uncertainty is best viewed in the frequency domain and describes errors that have bounded gain but arbitrary phase. Of course, there are types of uncertainty that do not fit succinctly into either of these two categories (e.g., state space uncertainty in which some time variation is allowed, or frequency domain uncertainty in which the phase is also bounded). Hence in practice, there are "shades of grey" when describing model uncertainty.

This research considers control design for nonparametric uncertainty. This type of uncertainty can be incorporated into the control design process using the small gain theorem. This theorem considers only one-block uncertainty. Unfortunately, for many systems the uncertainty occurs simultaneously in disparate parts. For example, in a model of a flexible structure, the errors might exist in the sensor and actuater dynamics in addition to errors which exist due to unmodeled dynamics. When uncertainty is present in the system in various places, control synthesis based solely on the small gain theorem may yield conservative control laws since the model of the uncertainty will then take into account errors that are not in the true uncertainty set. This conservatism motivated the development of the structured singular value.

The standard method for controller synthesis based on the structured singular value is usually referred to as "$D - K$ iteration." This process begins by fixing the D-scales defining an upper bound on the structured singular value (usually to $D = I$) and designing an H_∞ optimal controller K. Then with K fixed the D-scale magnitudes are optimized over (theoretically) all frequencies. Some optimal curve fit is then needed to find rational transfer functions that approximate the optimal D-scale magnitude plots (vs. frequency). Then, with the D-scales fixed to their rational transfer function approximations another H_∞ controller K is designed. The D-scales are then reoptimized (with K fixed). This process continues until convergence or until an acceptable controller is found.

Standard $D - K$ iteration with curve fitting has the advantage that at each iteration, a convex

optimization problem is solved, although the overall design process is *not* convex. However, this process also has serious drawbacks. First, there may not be a rational transfer function that corresponds to the optimal D-scale magnitude plot (vs. frequency). Even if such a function exists, it may be of very high order. If a low order transfer function is used, the design process will lead to a suboptimal controller. In fact, the resulting controller will generally not be the optimal controller for the D-scale of the given order.

This research develops a method for structured singular value controller synthesis that does not require curve fitting. In particular, the designer is allowed to *a priori* constrain the D-scales to be constant. The approach here is based on recent results in mixed norm H_2/H_∞ theory. As illustrated by Figure 12.1, for D-scales of a given order, the resultant controllers can have better robustness properties than those obtained using standard $D - K$ iteration and curve fitting.
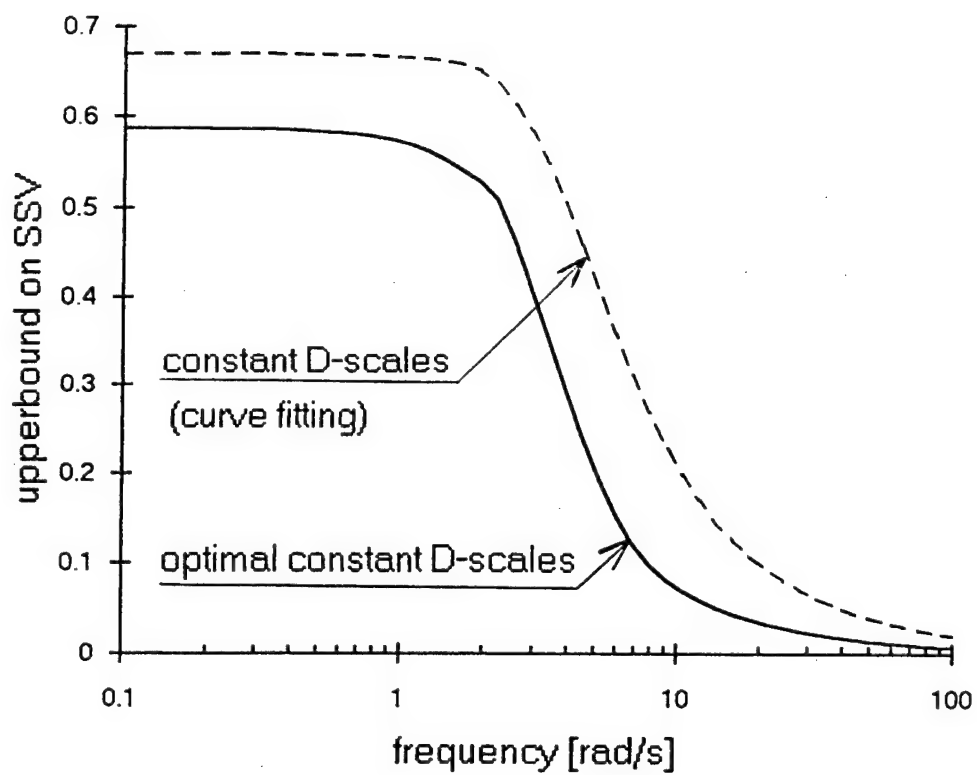
Figure 12.1. Upper Bounds on the CSSV Using the (Optimal)
Fixed Structure Approach and Standard D-K Iteration

Appendix A:

Maximum Entropy-Type Lyapunov Functions
for Robust Stability and Performance Analysis

Maximum-entropy-type Lyapunov functions for robust stability and performance analysis*

Dennis S. Bernstein

Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109-2140, USA

Wassim M. Haddad

Department of Mechanical and Aerospace Engineering, Florida Institute of Technology, Melbourne, FL 32901, USA

David C. Hyland

Harris Corporation, Government Aerospace Systems Division, MS 22/4847, Melbourne, FL 32902, USA

Feng Tyan

Department of Aerospace Engineering, The University of Michigan, Ann Arbor, MI 48109-2140, USA

Abstract: We present two Lyapunov functions that ensure the unconditional stability and robust performance of a modal system with uncertain damped natural frequency. Each Lyapunov function involves the sum of two matrices, the first being the solution to the so-called maximum-entropy equation and the second being a constant auxiliary portion. The significant feature of these Lyapunov functions is that the guaranteed robust stability region is independent of the weighting matrix, while the performance bounds are relatively tight compared to alternative approaches. Thus, these Lyapunov functions are less conservative than standard bounds that tend to be highly sensitive to the choice of state space basis.

Keywords: Maximum-entropy function; robust stability; robust performance

1. Introduction

The maximum-entropy approach to robust control was specifically developed to address the problem of modal uncertainty in flexible structures [2, 5, 6, 18, 19]. The rationale for this approach was based upon insights from the statistical analysis of lightly damped structures [20]. Despite favorable comparisons to other approaches [9, 10, 12, 13] and experimental application [11], the basis and meaning of the approach remain mostly empirical and largely obscure. The purpose of this paper is to make significant progress in developing a rigorous foundation for this approach.

Besides the statistical modal analysis techniques of [20], a variety of formulations have been put forth for justifying the maximum-entropy approach. To reproduce certain covariance phenomena of uncertain

multimodal systems (decorrelation, incoherence, and equipartition; see [20]), a multiplicative white-noise model was invoked [18, 19]. The specific model chosen was interpreted in the sense of Stratonovich, thus entailing a critical correction term in the covariance equation due to the conversion from Stratonovich to Ito calculus. The Stratonovich model was itself based upon a limiting process in which the parameter entropy increased, thus suggesting the name "maximum-entropy" control. White-noise models as a basis for robust control are discussed in [1].

An alternative justification for the maximum-entropy model was given in [14] in terms of positive real transfer functions. This attempt was motivated by the observation that in the limit of high modal frequency uncertainty the maximum-entropy controller assumed a rate dissipative structure [18, 19]. An alternative attempt to justify the maximum-entropy model was given in [17], where a covariance averaging approach [16] was used to show that if the state covariance is averaged over uncertain modal frequencies possessing a Cauchy distribution, then the resulting averaged covariance satisfies the maximum-entropy covariance model.

Although the various formulations of maximum-entropy theory lend considerable insight into the nature of the approach, there remains a significant gap between this approach and more conventional techniques, such as H_x theory. The missing link, in our opinion, is the lack of a Lyapunov function that guarantees the robust stability of the closed-loop control system. In this regard it was long suspected that such a Lyapunov function would be unconventional, that is, unlike those arising in H_x theory. This view arose from the fact that the maximum-entropy controllers were often robust to large perturbations in the damped natural frequencies, that is, the imaginary part of the eigenvalues. Such perturbations are highly structured, and thus are often treated conservatively by conventional small-gain-type bounds.

The goal of the present paper is to provide a Lyapunov function basis for the maximum-entropy covariance model for the case of modal frequency uncertainty. In fact, in this special case, we provide two alternative Lyapunov functions along with the corresponding performance bounds. Each Lyapunov function involves the sum of two matrices, the first being the solution to the maximum-entropy equation (see equation (22)) and the second being a constant auxiliary portion. This construction is similar to the parameter-dependent Lyapunov function technique developed in [15] except that in the present paper the auxiliary portion is constant, that is, independent of the uncertainty.

The maximum-entropy equation (22) differs fundamentally from alternative robustness tests such as those given in [3, 4]. Specifically, whereas the modified Lyapunov functions in [3] involve additional nonnegative-definite terms in the Lyapunov equation, the maximum-entropy equation entails an *indefinite* modification. This distinction appears to play a critical role with respect to the way in which the maximum-entropy equation deals with the change in basis induced by the input and weighting matrices.

While this paper potentially provides a Lyapunov function foundation for the maximum-entropy control approach, our results are limited to open-loop analysis. Future research will focus on robust stability of the closed-loop system for the controllers given in [2, 5, 6, 9–13, 18–20]. Furthermore, although the techniques used to construct the Lyapunov functions for the maximum-entropy equation are limited to modal frequency uncertainty, they appear to be generalizable to larger classes of uncertainty. Nevertheless, for structures with modal frequency uncertainty [2, 5, 6, 9–13, 18, 19], these results have practical ramifications.

2. Robust stability and performance problems

Let $\mathcal{U} \subset \mathbb{R}^{n \times n}$ denote a set of perturbations ΔA of a given nominal dynamics matrix $A \in \mathbb{R}^{n \times n}$. It is assumed that A is asymptotically stable and that $0 \in \mathcal{U}$.

Robust stability problem. Determine whether the linear system

$$\dot{x}(t) = (A + \Delta A)x(t), \quad t \in [0, \infty), \tag{1}$$

is asymptotically stable for all $\Delta A \in \mathcal{U}$.

Robust performance problem. For the disturbed linear system

$$\dot{x}(t) = (A + \Delta A)x(t) + Dw(t), \quad t \in [0, \infty), \tag{2}$$

$$z(t) = Ex(t), \tag{3}$$

where $w(\cdot)$ is a zero-mean d-dimensional white-noise signal with intensity I_d, determine a performance bound β satisfying

$$\mathscr{T}(\mathscr{U}) \triangleq \sup_{\Delta A \in \mathscr{U}} \limsup_{t \to \infty} E\{\|z(t)\|_2^2\} \leq \beta. \tag{4}$$

For convenience, define the $n \times n$ nonnegative-definite matrices $R \triangleq E^T E$ and $V \triangleq DD^T$. The following result is immediate. For a proof, see [3].

Lemma 2.1. *Suppose $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathscr{U}$. Then*

$$\mathscr{T}(\mathscr{U}) = \sup_{\Delta A \in \mathscr{U}} \operatorname{tr}(Q_{\Delta A} R) = \sup_{\Delta A \in \mathscr{U}} \operatorname{tr}(P_{\Delta A} V), \tag{5}$$

where $Q_{\Delta A} \in \mathbb{R}^{n \times n}$ and $P_{\Delta A} \in \mathbb{R}^{n \times n}$ are the unique, nonnegative-definite solutions to

$$0 = (A + \Delta A)Q_{\Delta A} + Q_{\Delta A}(A + \Delta A)^T + V \tag{6}$$

and

$$0 = (A + \Delta A)^T P_{\Delta A} + P_{\Delta A}(A + \Delta A) + R. \tag{7}$$

Conditions for robust stability and robust performance are developed in the following theorem. Let $\mathscr{N}^n$ and $\mathscr{S}^n$ denote the sets of $n \times n$ nonnegative-definite and symmetric matrices, respectively.

Theorem 2.2. *Let $\Omega_0: \mathscr{N}^n \to \mathscr{S}^n$, and suppose there exists $P \in \mathscr{N}^n$ satisfying*

$$0 = A^T P + PA + \Omega_0(P) + R. \tag{8}$$

Furthermore, let $P_0: \mathscr{U} \to \mathscr{S}^n$ and $R_0 \in \mathscr{S}^n$ be such that $R_0 \leq R$,

$$\Delta A^T P + P \Delta A \leq \Omega(P, \Delta A) + R_0, \quad \Delta A \in \mathscr{U}, \tag{9}$$

and

$$P + P_0(\Delta A) \geq 0, \quad \Delta A \in \mathscr{U}, \tag{10}$$

where

$$\Omega(P, \Delta A) \triangleq \Omega_0(P) - [(A + \Delta A)^T P_0(\Delta A) + P_0(\Delta A)(A + \Delta A)]. \tag{11}$$

Then

$$(R - R_0, A + \Delta A), \quad \Delta A \in \mathscr{U}, \tag{12}$$

is detectable if and only if

$$A + \Delta A, \quad \Delta A \in \mathscr{U}, \tag{13}$$

is asymptotically stable. In this case, the following statements are true. If $\gamma < 1$ is such that $R_0 \leq \gamma R$, then

$$P_{\Delta A} \leq \frac{1}{1 - \gamma}(P + P_0(\Delta A)), \quad \Delta A \in \mathscr{U}, \tag{14}$$

where $P_{\Delta A}$ satisfies (7), and

$$\mathscr{T}(\mathscr{U}) \leq \frac{1}{1 - \gamma}[\operatorname{tr}(PV) + \sup_{\Delta A \in \mathscr{U}} \operatorname{tr}(P_0(\Delta A)V)]. \tag{15}$$

In addition, if there exists $\bar{P}_0 \in \mathscr{S}^n$ such that

$$P_0(\Delta A) \le \bar{P}_0,\tag{16}$$

then

$$\mathscr{T}(\mathscr{U}) \le \frac{1}{1-\gamma}\,\mathrm{tr}[(P+\bar{P}_0)V].\tag{17}$$

Proof. Note that, for all $\Delta A \in \mathscr{U}$, (8) is equivalent to

$$
\begin{aligned}
0 &= (A+\Delta A)^{\mathrm{T}}(P+P_0(\Delta A)) + (P+P_0(\Delta A))(A+\Delta A) + \Omega_0(P) + R \\
&\quad - [(A+\Delta A)^{\mathrm{T}}P_0(\Delta A) + P_0(\Delta A)(A+\Delta A)] - (\Delta A^{\mathrm{T}}P + P\Delta A) \\
&= (A+\Delta A)^{\mathrm{T}}(P+P_0(\Delta A)) + (P+P_0(\Delta A))(A+\Delta A) + R - R_0 + R_0',
\end{aligned}\tag{18}
$$

where

$$
\begin{aligned}
R_0' &\triangleq \Omega_0(P) + R_0 - [(A+\Delta A)^{\mathrm{T}}P_0(\Delta A) + P_0(\Delta A)(A+\Delta A)] - (\Delta A^{\mathrm{T}}P + P\Delta A) \\
&= \Omega(P,\Delta A) + R_0 - (\Delta A^{\mathrm{T}}P + P\Delta A).
\end{aligned}
$$

Hence, (18) has a solution $P \in \mathscr{N}^n$ for all $\Delta A \in \mathscr{U}$. Thus, if the detectability condition (12) holds for all $\Delta A \in \mathscr{U}$, then it follows from [21, Theorem 3.6] that $(R - R_0 + R_0', A + \Delta A)$ is detectable, $\Delta A \in \mathscr{U}$. It now follows from (18) and [21, Lemma 12.2] that $A + \Delta A$ is asymptotically stable, $\Delta A \in \mathscr{U}$. Conversely, if $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathscr{U}$, then (12) is immediate.

Now, subtracting $(1-\gamma)\cdot(7)$ from (18) yields

$$
\begin{aligned}
0 &= (A+\Delta A)^{\mathrm{T}}(P+P_0(\Delta A) - (1-\gamma)P_{\Delta A}) + (P+P_0(\Delta A) - (1-\gamma)P_{\Delta A})(A+\Delta A) \\
&\quad + R_0' - R_0 + \gamma R, \quad \Delta A \in \mathscr{U},
\end{aligned}\tag{19}
$$

or, since $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathscr{U}$ and $R_0 \le \gamma R$, (19) implies that, for all $\Delta A \in \mathscr{U}$,

$$
\begin{aligned}
P + P_0(\Delta A) - (1-\gamma)P_{\Delta A} &= \int_0^\infty e^{(A+\Delta A)^{\mathrm{T}}t}[R_0' + \gamma R - R_0]e^{(A+\Delta A)t}\,dt \\
&\ge \int_0^\infty e^{(A+\Delta A)^{\mathrm{T}}t}R_0' e^{(A+\Delta A)t}\,dt \\
&\ge 0,
\end{aligned}
$$

which implies (14).

Next, using (14), it follows from (5) that

$$
\begin{aligned}
\mathscr{T}(\mathscr{U}) = \sup_{\Delta A \in \mathscr{U}} \mathrm{tr}(D^{\mathrm{T}}P_{\Delta A}D) &\le \frac{1}{1-\gamma}\sup_{\Delta A \in \mathscr{U}} \mathrm{tr}[D^{\mathrm{T}}(P+P_0(\Delta A))D] \\
&= \frac{1}{1-\gamma}\left[\mathrm{tr}(PV) + \sup_{\Delta A \in \mathscr{U}} \mathrm{tr}(P_0(\Delta A)V)\right],
\end{aligned}
$$

which yields (15). Furthermore, using (16) it follows that

$$
\mathscr{T}(\mathscr{U}) \le \frac{1}{1-\gamma}\left[\mathrm{tr}(PV) + \sup_{\Delta A \in \mathscr{U}} \mathrm{tr}(P_0(\Delta A)V)\right] \le \frac{1}{1-\gamma}[\mathrm{tr}(PV) + \mathrm{tr}(\bar{P}_0 V)]
$$

$$
= \frac{1}{1-\gamma}\,\mathrm{tr}[(P+\bar{P}_0)V]. \qquad \square
$$

Remark 2.3. Theorem 2.2 is a generalization of Theorem 3.1 of [15]. Specifically, the bound in [15] is required to hold for all nonnegative-definite matrices, whereas in Theorem 2.2 equation (9) need only hold for the solution P of (8). Furthermore, in [15], $R_0 = 0$.

Remark 2.4. Inequality (9) is equivalent to

$$(A + \Delta A)^{\mathrm{T}}(P + P_0(\Delta A)) + (P + P_0(\Delta A))(A + \Delta A) + R - R_0 \leq 0,$$

which shows that $V(x) = x^{\mathrm{T}}(P + P_0(\Delta A))x$ is a Lyapunov function corresponding to $A + \Delta A$. In constructing this Lyapunov function, the matrix P can be viewed as a *predictor* term, $P_0(\Delta A)$ provides a *corrector* term, and $P_{\mathrm{T}} \triangleq P + P_0(\Delta A)$ is the *total* Lyapunov matrix.

Remark 2.5. If $P_0(\Delta A)$ is independent of ΔA, then by choosing $\bar{P}_0 = P_0(\Delta A)$ it follows that (15) is identical to (17).

3. Application to the maximum-entropy covariance model

Now we specialize to the case in which $\mathcal{U}$ is given by

$$\mathcal{U} \triangleq \left\{ \Delta A \in \mathbb{R}^{n \times n} : \Delta A = \sum_{i=1}^{r} \sigma_i A_i, \ |\sigma_i| \leq \delta_i, \quad i = 1, \ldots, r \right\}, \tag{20}$$

where $\delta_i > 0$ and the matrices $A_i \in \mathbb{R}^{n \times n}$, which represent the uncertainty structure, are the given skew-symmetric matrices, that is, $A_i + A_i^{\mathrm{T}} = 0$, $i = 1, \ldots, r$. In addition, we assume that $A + A^{\mathrm{T}} < 0$. This formulation can be viewed as the representation of a dissipative system (such as a flexible structure) with energy-conserving perturbations. This property can be seen by means of the Lyapunov function $V(x) = x^{\mathrm{T}}x$ whose decay rate is independent of σ_i. Thus, $A + \Delta A$ is uniformly asymptotically stable even for arbitrarily time-varying $\sigma_i(t)$. For simplicity, however, we confine our analysis to constant parameter uncertainty. In addition, although the system is robustly stable for time-varying parameter uncertainties, the performance bounds we obtain via Theorem 2.2 are valid only for the case of constant parameter uncertainty.

We now introduce a specific choice of $\Omega_0(P)$ that is motivated by the maximum-entropy covariance model. Specifically, as in [18] we choose

$$\Omega_0(P) = \sum_{i=1}^{r} \delta_i^2 (\tfrac{1}{2} A_i^{2\mathrm{T}} P + A_i^{\mathrm{T}} P A_i + \tfrac{1}{2} P A_i^2). \tag{21}$$

First we prove that with this choice of $\Omega_0(P)$ equation (8) has a unique solution. Then we show that, when $r = 1$, equation (8) has an asymptotic solution for $\delta_1 \to \infty$.

Proposition 3.1. *Assume that $A + A^{\mathrm{T}} < 0$, $A_i + A_i^{\mathrm{T}} = 0$, and $\delta_i \geq 0$, $i = 1, \ldots, r$. Then there exists a unique matrix $P \in \mathbb{R}^{n \times n}$ satisfying*

$$0 = A^{\mathrm{T}} P + PA + \sum_{i=1}^{r} \delta_i^2 (\tfrac{1}{2} A_i^{2\mathrm{T}} P + A_i^{\mathrm{T}} P A_i + \tfrac{1}{2} P A_i^2) + R. \tag{22}$$

Furthermore, P is nonnegative-definite.

Proof. Applying the "vec" operator [7] to (22) yields

$$0 = \mathcal{A}^{\mathrm{T}} \operatorname{vec} P + \operatorname{vec} R, \tag{23}$$

where

$$\mathcal{A} \triangleq (A \oplus A) + \sum_{i=1}^{r} \tfrac{1}{2} \delta_i^2 (A_i \oplus A_i)^2$$

and $\oplus$ and later $\otimes$ denote Kronecker sum and product, respectively. Since $A + A^T < 0$, it follows that $(A \oplus A) + (A \oplus A)^T = (A + A^T) \oplus (A + A^T) < 0$. In addition, the assumption that A_i is skew-symmetric implies that $A_i \oplus A_i$ is also skew-symmetric and thus $(A_i \oplus A_i)^2 \leq 0$, $i = 1, \ldots, r$. Thus, $\mathscr{A} + \mathscr{A}^T < 0$, which implies that $\mathscr{A}$ is asymptotically stable. Thus, (23) yields $P = \text{vec}^{-1}(-\mathscr{A}^{-T} \text{vec } R)$. This proves existence and uniqueness.

Next, we show that P is nonnegative-definite. Note that since $-\mathscr{A}^{-T} = \int_0^\infty e^{\mathscr{A}^T t} \, dt$, we can write

$$P = \text{vec}^{-1}\left(\int_0^\infty e^{\mathscr{A}^T t} \text{vec } R \, dt\right). \tag{24}$$

After some manipulation (24) can be written as

$$P = \text{vec}^{-1}\left(\int_0^\infty \exp\left(t\left[\sum_{i=1}^r \left(\frac{A}{r} + \tfrac{1}{2}\delta_i^2 A_i^2\right)^T \oplus \sum_{i=1}^r \left(\frac{A}{r} + \tfrac{1}{2}\delta_i^2 A_i^2\right)^T + \sum_{i=1}^r \tfrac{1}{2}\delta_i^2 (A_i \otimes A_i)^T\right]\right) \text{vec } R \, dt\right). \tag{25}$$

Now, using the exponential product formula it follows that

$$P = \text{vec}^{-1}\left(\int_0^\infty \lim_{m \to \infty} \left[\exp\left(\frac{t}{n}\left[\sum_{i=1}^r \left(\frac{A^T}{r} + \tfrac{1}{2}\delta_i^2 A_i^{2T}\right) \oplus \sum_{i=1}^r \left(\frac{A^T}{r} + \tfrac{1}{2}\delta_i^2 A_i^{2T}\right)\right)\right)\right]\right)$$

$$\times \prod_{i=1}^r \exp\left(\frac{\delta_i^2 t}{2m}(A_i^T \otimes A_i^T)\right)\bigg]^m \text{vec } R \, dt\bigg). \tag{26}$$

For simplicity, we assume $r = 1$. If $r > 1$ only minor modifications are needed. First fix m and let $R_{(0)} \triangleq R$; define the series $Z_{(j)}, R_{(j)}, j = 0, 1, \ldots, m - 1$, by

$$\text{vec } Z_{(j+1)}(t) \triangleq e^{(\delta_1^2 t/2m)(A_1 \otimes A_1)^T} \text{vec } R_{(j)}(t) = \text{vec } \sum_{k=0}^\infty \frac{1}{k!}\left(\frac{\delta_1^2 t}{2m}\right)^k A_1^{kT} R_{(j)}(t) A_1^k,$$

$$\text{vec } R_{(j+1)}(t) \triangleq \exp\left(\frac{t}{m}\left(A + \frac{\delta_1^2}{2}A_1^2\right) \oplus \left(A + \frac{\delta_1^2}{2}A_1^2\right)^T\right) \text{vec } Z_{(j+1)}(t)$$

$$= \text{vec } \exp\left(\frac{t}{m}\left(A + \frac{\delta_1^2}{2}A_1^2\right)^T\right) Z_{(j+1)}(t) \exp\left(\frac{t}{m}\left(A + \frac{\delta_1^2}{2}A_1^2\right)\right).$$

It is obvious that both $Z_{(j)}(t)$ and $R_{(j)}(t)$ are nonnegative-definite matrices for all $j = 0, 1, \ldots, m - 1$ and $t \geq 0$. Finally, since m is arbitrary, it can be shown that

$$P = \text{vec}^{-1}\left(\int_0^\infty \lim_{m \to \infty} \text{vec } R_{(m)} \, dt\right) = \int_0^\infty \lim_{m \to \infty} R_{(m)} \, dt \geq 0. \quad \square$$

Next we show that (22) with $r = 1$ has an asymptotic solution for $\delta_1 \to \infty$. First, we need the following definition and lemma.

Definition 3.2. For $F \in \mathbb{R}^{n \times n}$, the smallest nonnegative integer k such that rank $(F^k) = $ rank (F^{k+1}) is called the *index* of F and is denoted by Ind (F) [8].

Remark 3.3. If F is invertible, Ind $(F) = 0$. Also Ind $(0) = 1$. We adopt the convention that $0^0 = 1$ [8].

Definition 3.4. A matrix $F \in \mathbb{R}^{n \times n}$ is called EP [8] if either F is invertible or there exists an orthogonal matrix $U \in \mathbb{R}^{n \times n}$ and an invertible matrix $F_1 \in \mathbb{R}^{m \times m}$, where $m \leq n$, such that

$$F = U \begin{bmatrix} F_1 & 0 \\ 0 & 0 \end{bmatrix} U^T.$$

Remark 3.5. If F is EP, then $\mathrm{Ind}\,(F) \leq 1$, and the group inverse F^* of F is given by [8]

$$F^* = U \begin{bmatrix} F_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^{\mathrm{T}}.$$

Lemma 3.6. Let $A, B \in \mathbb{R}^{n \times n}$, where $A + A^{\mathrm{T}} < 0$ and B is an EP matrix. Then

$$\mathrm{Ind}\,(AB) = \mathrm{Ind}\,(B). \tag{27}$$

Proof. Since B is an EP matrix, Remark 3.5 implies that $\mathrm{Ind}\,(B) \leq 1$. Hence, we consider two cases.

(1) Suppose $\mathrm{Ind}\,(B) = 0$, so that B is invertible. Since $A + A^{\mathrm{T}} < 0$, it follows that A is asymptotically stable and hence invertible. Therefore, AB is invertible and thus $\mathrm{Ind}\,(AB) = 0$.

(2) Suppose $\mathrm{Ind}\,(B) = 1$, and let $\mathrm{rank}\,(B) = n - r$, where $r \geq 1$. Since B is an EP matrix, there exists an orthogonal matrix U and a matrix D_B such that $B = UD_BU^{\mathrm{T}}$, where

$$D_B = \begin{bmatrix} B_1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_1 \in \mathbb{R}^{(n-r) \times (n-r)}, \det\,(B_1) \neq 0.$$

Since $\mathrm{rank}\,(AB) = n - r$, it suffices to show that the zero eigenvalue of AB has multiplicity r.

By writing $U^{\mathrm{T}}AU$ in the form

$$A' \triangleq U^{\mathrm{T}}AU = \begin{bmatrix} A'_{11} & A'_{12} \\ A'_{21} & A'_{22} \end{bmatrix},$$

where $A'_{11} \in \mathbb{R}^{(n-r) \times (n-r)}$, $A'_{22} \in \mathbb{R}^{r \times r}$, $A'_{12} \in \mathbb{R}^{(n-r) \times r}$, $A'_{21} \in \mathbb{R}^{r \times (n-r)}$, we have

$$U^{\mathrm{T}}AUD_B = \begin{bmatrix} A'_{11}B_1 & 0 \\ A'_{21}B_1 & 0 \end{bmatrix}.$$

Consequently, the characteristic polynomial of AB is

$$\det\,(\lambda I - AB) = \det\,(\lambda I - U(U^{\mathrm{T}}AU D_B)U^{\mathrm{T}}) = \det\,(\lambda I - U^{\mathrm{T}}AU D_B)$$

$$= \det \begin{bmatrix} \lambda I_{n-r} - A'_{11}B_1 & 0 \\ -A'_{21}B_1 & \lambda I_r \end{bmatrix} = \lambda^r \det\,(\lambda I_{n-r} - A'_{11}B_1). \tag{28}$$

Equation (28) implies that the zero eigenvalue of AB has at least multiplicity r.

The final step is to show that $A'_{11}B_1$ has no zero eigenvalue or, equivalently, $\det\,(A'_{11}B_1) \neq 0$. Since $A + A^{\mathrm{T}} < 0$, it follows that $U^{\mathrm{T}}(A + A^{\mathrm{T}})U < 0$, that is, $A' + A'^{\mathrm{T}} < 0$. Thus, $A'_{11} + (A'_{11})^{\mathrm{T}} < 0$, which implies that A'_{11} is asymptotically stable. Therefore, we have $\det\,(A'_{11}) \neq 0$. Noting

$$\det\,(A'_{11}B_1) = \det\,(A'_{11})\det\,(B_1) \neq 0$$

completes the proof. $\square$

For convenience, we define

$$\Lambda \triangleq (A^{\mathrm{T}} \oplus A^{\mathrm{T}})^{-1}(A_1^{\mathrm{T}} \oplus A_1^{\mathrm{T}})^2. \tag{29}$$

Lemma 3.7. Let $A, A_1 \in \mathbb{R}^{n \times n}$, where $A + A^{\mathrm{T}} < 0$ and $A_1 + A_1^{\mathrm{T}} = 0$. Then $\mathrm{Ind}\,(\Lambda) = 1$.

Proof. Since A_1 is skew-symmetric, it follows that $A_1 \oplus A_1$ is also skew-symmetric. Thus, $(A_1 \oplus A_1)^2$ is symmetric (actually, it is negative-semidefinite) and hence is EP. In addition, it is obvious that $A_1 \oplus A_1$ is singular. Thus, $\mathrm{Ind}\,(A_1^{\mathrm{T}} \oplus A_1^{\mathrm{T}})^2 = 1$. Furthermore, since $A + A^{\mathrm{T}} < 0$ implies $(A \oplus A) + (A^{\mathrm{T}} \oplus A^{\mathrm{T}}) < 0$ and equivalently implies $(A \oplus A)^{-1} + (A^{\mathrm{T}} \oplus A^{\mathrm{T}})^{-1} < 0$, it follows from Lemma 3.6 that $\mathrm{Ind}\,(\Lambda) = 1$. $\square$

We are now ready to prove the existence of an asymptotic solution of equation (8) when $r = 1$. For notational convenience, we replace $\delta_1^2/2$ by α.

Proposition 3.8. *Let $A, A_1 \in \mathbb{R}^{n \times n}$, $R \in \mathcal{N}^n$ and $\alpha \geq 0$. Furthermore, assume that $A + A^T < 0$, $A_1 + A_1^T = 0$, and let $P_\alpha \in \mathcal{N}^n$ be the unique, nonnegative-definite solution to*

$$0 = A^T P + PA + \alpha(A_1^{2T}P + 2A_1^T PA_1 + PA_1^2) + R. \tag{30}$$

Then $P_\infty \triangleq \lim_{\alpha \to \infty} P_\alpha$ exists and is given by

$$P_\infty = \text{vec}^{-1}[(I - \Lambda\Lambda^*)(A^T \oplus A^T)^{-1}(-\text{vec } R)]. \tag{31}$$

Proof. Applying the vec operator to equation (30) yields

$$0 = [(A^T \oplus A^T) + \alpha(A_1^T \oplus A_1^T)^2]\text{vec } P + \text{vec } R,$$

so that

$$\text{vec } P = [I + \alpha\Lambda]^{-1}(A^T \oplus A^T)^{-1}(-\text{vec } R),$$

and we can write P_∞ as

$$\text{vec } P_\infty = \lim_{\alpha \to \infty}(I + \alpha\Lambda)^{-1}(A^T \oplus A^T)^{-1}(-\text{vec } R)$$

$$= \lim_{\alpha \to \infty}\left[\alpha\left(\frac{1}{\alpha}I + \Lambda\right)\right]^{-1}(A^T \oplus A^T)^{-1}(-\text{vec } R)$$

$$= \lim_{z \to \infty} z(zI + \Lambda)^{-1}(A^T \oplus A^T)^{-1}(-\text{vec } R).$$

Now since $\text{Ind}(\Lambda) = 1$, it follows from [8, Theorem 7.6.2] that the above limit exists and is given by $\text{vec } P_\infty = (I - \Lambda\Lambda^*)(A^T \oplus A^T)^{-1}(-\text{vec } R)$, which yields (31). $\square$

For the following result, define the commutator $[F, G] \triangleq FG - GF$.

Lemma 3.9. *Let $A, A_1 \in \mathbb{R}^{n \times n}$, $R \in \mathcal{N}^n$. Furthermore, suppose that $A + A^T < 0$, $A_1 + A_1^T = 0$, and let $P_\infty \in \mathcal{N}^n$ be given by (31). Then P_∞ satisfies*

$$[A_1^T, P_\infty] = 0. \tag{32}$$

Proof. Since A_1 is skew-symmetric, we have

$$\text{vec}[A_1^T, P_\infty] = \text{vec}(A_1^T P_\infty + P_\infty A_1) = (A_1^T \oplus A_1^T)\text{vec } P_\infty$$

$$= (A_1^T \oplus A_1^T)(I - \Lambda\Lambda^*)(A^T \oplus A^T)^{-1}(-\text{vec } R), \tag{33}$$

where Λ is defined by (29). Since, by Lemma 3.7, $\text{Ind}(\Lambda) = 1$, it follows from Remark 3.5 that Λ and Λ^* can be expressed in the form

$$\Lambda = V\begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix}V^{-1}, \qquad \Lambda^* = V\begin{bmatrix} C^{-1} & 0 \\ 0 & 0 \end{bmatrix}V^{-1},$$

where $\det(C) \neq 0$. Writing $V = [V_1 \ V_2]$, the identity

$$\Lambda V = V\begin{bmatrix} C & 0 \\ 0 & 0 \end{bmatrix}$$

implies that $\Lambda V_2 = 0$. Consequently, $(A_1^T \oplus A_1^T)^2 V_2 = 0$, and, since $\mathrm{Ind}\,(A_1^T \oplus A_1^T) = 1$, it follows that $(A_1^T \oplus A_1^T) V_2 = 0$. Therefore, equation (33) can be written as

$$
\begin{aligned}
\mathrm{vec}\,[A_1^T, P_x] &= (A_1^T \oplus A_1^T)\left(I - V\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} V^{-1}\right)(A^T \oplus A^T)^{-1}(-\,\mathrm{vec}\,R) \\
&= (A_1^T \oplus A_1^T)\left(V\begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} V^{-1}\right)(A^T \oplus A^T)^{-1}(-\,\mathrm{vec}\,R) \\
&= (A_1^T \oplus A_1^T)\,[0 \;\, V_2]\,V^{-1}(A^T \oplus A^T)^{-1}(-\,\mathrm{vec}\,R) \\
&= [0\,(A_1^T \oplus A_1^T)\,V_2]\,V^{-1}(A^T \oplus A^T)^{-1}(-\,\mathrm{vec}\,R) = 0.
\end{aligned}
$$

As a result, $[A_1^T, P_x] = 0$. $\quad\square$

Remark 3.10. If P is symmetric, A_1 is skew-symmetric, then it can be shown that $[A_1^T, [A_1^T, P_x]] = 0$ if and only if $[A_1^T, P_x] = 0$. This fact is of interest since (21) can be written as

$$
\Omega_0(P) = \sum_{i=1}^{r} \tfrac{1}{2}\delta_i^2\,[A_i^T, [A_i^T, P]].
$$

Thus, if $r = 1$ and $\delta_1 \to \infty$, then $[A_1^T, [A_1^T, P_\infty]] \to 0$. Note $(\delta_1^2/2)\,[A_1^T, [A_1^T, P_\infty]] = -(A^T P_\infty + P_\infty A + R) = -\,\mathrm{vec}^{-1}[(A_1^T \oplus A_1^T)^2[(A^T \oplus A^T)^{-1}(A_1^T \oplus A_1^T)^2]^* (A^T \oplus A^T)^{-1}\,\mathrm{vec}\,R]$.

4. The choice of corrector term P_0

Now we propose a corrector term P_0 for the case of general skew-symmetric matrices $A_i \in \mathbb{R}^{n \times n}$, $i = 1, \dots, r$, where $r \geq 1$. For a symmetric matrix B, define $|B| \triangleq \sqrt{B^2}$.

Proposition 4.1. Assume $A + A^T < 0$, $A_i + A_i^T = 0$, and $\delta_i \geq 0$, $i = 1, \dots, r$. Let $P \in \mathcal{N}^n$ satisfy (22) and let

$$
\beta \geq \max\left\{\sum_{i=1}^{r} \mu_i,\; -\lambda_{\min}(P)\right\}, \tag{34}
$$

where, for $i = 1, \dots, r$,

$$
\mu_i \triangleq \lambda_{\max}((\delta_i|[A_i^T, P]| - \tfrac{1}{2}\delta_i^2[A_i^T, [A_i^T, P]])(-A - A^T)^{-1}).
$$

If $P_0(\Delta A) \triangleq \beta I_n$, then (9) and (10) are satisfied with $R_0 = 0$ and $\mathcal{U}$ given by (20).

Proof. By substituting $P_0(\Delta A) = \beta I_n$ into (9) with $R_0 = 0$ and letting $G = \sqrt{-A^T - A}$, we have

$$
\Omega(P, \Delta A) + R_0 - (\Delta A^T P + P \Delta A)
$$

$$
= \beta(-A^T - A) - \sum_{i=1}^{r} \sigma_i[A_i^T, P] + \sum_{i=1}^{r} \tfrac{1}{2}\delta_i^2[A_i^T, [A_i^T, P]]
$$

$$
\geq \beta(-A^T - A) - \sum_{i=1}^{r} \delta_i|[A_i^T, P]| + \sum_{i=1}^{r} \tfrac{1}{2}\delta_i^2[A_i^T, [A_i^T, P]]
$$

$$
= G\left\{\beta I_n - \sum_{i=1}^{r} G^{-1}(\delta_i|[A_i^T, P]| - \tfrac{1}{2}\delta_i^2[A_i^T, [A_i^T, P]])G^{-1}\right\}G
$$

$$
\geq G\left\{\beta I_n - \sum_{i=1}^{r} \lambda_{\max}(G^{-1}(\delta_i|[A_i^T, P]| - \tfrac{1}{2}\delta_i^2[A_i^T, [A_i^T, P]])G^{-1})I_n\right\}G
$$

$$= G\{\beta I_n - \sum_{i=1}^{r} \lambda_{\max}((\delta_i |[A_i^\mathsf{T}, P]| - \tfrac{1}{2}\delta_i^2 [A_i^\mathsf{T}, [A_i^\mathsf{T}, P]])(-A - A^\mathsf{T})^{-1})I_n\}G$$

$$= G\{\beta I_n - \sum_{i=1}^{r} \mu_i I_n\}G$$

$$\geq 0,$$

which proves (9). Finally, it is obvious that $P + P_0(\Delta A) = P + \beta I_n \geq \lambda_{\min}(P) I_n + \beta I_n \geq 0$, so that (10) is satisfied. $\square$

Henceforth, we confine our attention to the special case $r = 1$ and

$$A = \begin{bmatrix} -\eta & \omega \\ -\omega & -\eta \end{bmatrix}, \qquad A_1 = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \tag{35}$$

where $\eta > 0$ and $\omega \in \mathbb{R}$. For notational convenience, we adopt the traditional symbol J for A_1. In this case $\Omega_0(P)$ given by (21) has the form

$$\Omega_0(P) = \delta_1^2 (\tfrac{1}{2} J^{2\mathsf{T}} P + J^\mathsf{T} P J + \tfrac{1}{2} P J^2). \tag{36}$$

Note that $J^\mathsf{T} = -J$ and $J^2 = -I_2$, where I_2 denotes the 2×2 identity matrix.

Proposition 4.2. *Assume that R is positive-definite and let P satisfy*

$$0 = A^\mathsf{T} P + PA + \delta_1^2 (\tfrac{1}{2} J^{2\mathsf{T}} P + J^\mathsf{T} P J + \tfrac{1}{2} P J^2) + R, \tag{37}$$

let $\gamma < 1$, and define

$$P_0(\Delta A) \triangleq (1 - \gamma) J^\mathsf{T} P J - \gamma P, \quad \Delta A \in \mathcal{U}. \tag{38}$$

Then (9) and (10) are satisfied with $R_0 = \gamma R$. Furthermore, the performance bound (15) is given by

$$\mathcal{T}(\mathcal{U}) \leq \operatorname{tr}(V) \operatorname{tr}(P). \tag{39}$$

Proof. Clearly, (10) is satisfied. Secondly, since

$$AJ = JA, \quad JJ^\mathsf{T} = J^\mathsf{T} J = I_2, \quad J^\mathsf{T} \Omega_0(P) J = -\Omega_0(P),$$

and P satisfies (37), it follows that

$$\Omega_0(P) + R_0 - [(A + \sigma_1 J)^\mathsf{T} P_0 + P_0(A + \sigma_1 J)] - \sigma_1(J^\mathsf{T} P + PJ)$$

$$= \Omega_0(P) + R_0 - [(1 - \gamma)(A^\mathsf{T} J^\mathsf{T} PJ + J^\mathsf{T} PJA) + \sigma_1(1 - \gamma)(J^\mathsf{T} J^\mathsf{T} PJ + J^\mathsf{T} PJJ)$$

$$\quad - \gamma(A^\mathsf{T} P + PA) - \sigma_1 \gamma(J^\mathsf{T} P + PJ)] - \sigma_1(J^\mathsf{T} P + PJ)$$

$$= \Omega_0(P) + R_0 - (1 - \gamma)J^\mathsf{T}(A^\mathsf{T} P + PA)J + \gamma(A^\mathsf{T} P + PA)$$

$$= \Omega_0(P) + R_0 - (1 - \gamma)J^\mathsf{T}(-\Omega_0(P) - R)J + \gamma(-\Omega_0(P) - R)$$

$$= R_0 - \gamma R + (1 - \gamma)J^\mathsf{T} RJ$$

$$\geq 0.$$

Finally, we have

$$\mathcal{T}(\mathcal{U}) \leq \frac{1}{1 - \gamma}[\operatorname{tr}(PV) + \operatorname{tr}(P_0 V)] = \operatorname{tr}(PV) + \operatorname{tr}(J^\mathsf{T} PJV)$$

$$= \operatorname{tr}[P(V + JVJ^\mathsf{T})] = \operatorname{tr}(V) \operatorname{tr}(P). \quad \square$$

Remark 4.3. Note that unlike the parameter-dependent Lyapunov function used in [15] for the Popov criterion, the auxiliary portion $P_0(\Delta A)$ given by (38) is independent of σ_1. Therefore, this auxiliary portion $P_0(\Delta A)$ guarantees robust stability with respect to time-varying $\sigma_1(t)$. This robust stability property was already shown at the beginning of this section by means of the Lyapunov function $V(x) = x^\mathsf{T} x$.

Remark 4.4. Since by Proposition 3.1, equation (37) has a solution for all $\delta_1 > 0$, it follows that robust stability is guaranteed for arbitrary σ_1, that is, not necessarily bounded by δ_1.

Remark 4.5. It is easy to show that $\operatorname{tr}(P) = (1/2\eta)\operatorname{tr}(R)$ and $P_T = P + P_0 = (1 - \gamma)(J^\mathsf{T} P J + P) = (1 - \gamma)\operatorname{tr}(P) I_2$. Thus, (39) becomes

$$\mathscr{T}(\mathscr{U}) \le \frac{1}{2\eta}\operatorname{tr}(V)\operatorname{tr}(R). \tag{40}$$

Thus, the performance bound (39) is independent of δ_1. Furthermore, it is easy to check that P_T satisfies the equation

$$0 = A^\mathsf{T} P_T + P_T A + J^\mathsf{T} R J + R. \tag{41}$$

We now present an alternative choice of $P_0(\Delta A)$.

Proposition 4.6. *Let*

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{bmatrix}, \qquad R = \begin{bmatrix} R_{11} & R_{12} \\ R_{12} & R_{22} \end{bmatrix} > 0$$

satisfy (37) *and let* $P_0(\Delta A) \triangleq \mu I_2$, *where*

$$\mu \triangleq \frac{\sqrt{\delta_1^2 + \delta_1^4}}{2\eta}\sqrt{(P_{22} - P_{11})^2 + (2P_{12})^2}. \tag{42}$$

Then (9) *and* (10) *are satisfied with* $R_0 = 0$. *Furthermore, the performance bound* (15) *is given by*

$$\mathscr{T}(\mathscr{U}) \le \operatorname{tr}(PV) + \mu\operatorname{tr}(V). \tag{43}$$

Proof. Since $P \ge 0$ and $P_0(\Delta A) \ge 0$, $\Delta A \in \mathscr{U}$, it follows that (10) is satisfied. Next, to show that (9) is true, recall that $\Omega_0(P)$ is given by equation (36). Therefore,

$$\Omega_0(P) - [(A + \sigma_1 J)^\mathsf{T} P_0 + P_0(A + \sigma_1 J)] - \sigma_1(J^\mathsf{T} P + PJ)$$

$$= \delta_1^2(-P + J^\mathsf{T} PJ) - \mu(A^\mathsf{T} + A) - \sigma_1(J^\mathsf{T} P + PJ)$$

$$= 2\mu\eta I_2 + \delta_1^2(-P + J^\mathsf{T} PJ) - \sigma_1(J^\mathsf{T} P + PJ)$$

$$= 2\mu\eta I_2 + S\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}S^\mathsf{T}$$

$$= S\begin{bmatrix} 2\mu\eta + \lambda_1 & 0 \\ 0 & 2\mu\eta + \lambda_2 \end{bmatrix}S^\mathsf{T},$$

where $\lambda_1 = -\lambda_2 = \sqrt{\sigma_1^2 + \delta_1^4}\sqrt{(P_{22} - P_{11})^2 + (2P_{12})^2}$ are the eigenvalues of $\delta_1^2(-P + J^\mathsf{T} PJ) - \sigma_1(J^\mathsf{T} P + PJ)$ and S is a 2×2 orthogonal matrix. Choosing μ according to (42) implies that $2\mu\eta + \lambda_1 \ge 0$ and $2\mu\eta + \lambda_2 \ge 0$. Thus, (9) is satisfied. Finally, the performance bound (15) has the form

$$\mathscr{T}(\mathscr{U}) \le \operatorname{tr}[(P + P_0(\Delta A))V] = \operatorname{tr}(PV) + \mu\operatorname{tr}(V). \qquad \square$$

Remark 4.7. As in [3, 4] the robust performance bounds (40) and (43) are only valid for constant uncertainty σ_1.

Before we present a numerical example, we shall illustrate some important aspects of P given by equation (37). The analytical solution for (37) yields

$$P_{11} + P_{22} = \frac{1}{2\eta}(R_{11} + R_{22}), \qquad P_{11} - P_{22} = \frac{1}{\alpha}\left[\frac{\eta + \delta_1^2}{2}(R_{11} - R_{22}) - \omega R_{12}\right],$$

$$2P_{12} = \frac{1}{\alpha}\left[\frac{\omega}{2}(R_{11} - R_{22}) + (\eta + \delta_1^2)R_{12}\right],$$

where $\alpha \triangleq (\eta + \delta_1^2)^2 + \omega^2$. For large δ_1, it is easy to see that

$$P_{11} - P_{22} \sim \frac{1}{2\delta_1^2}(R_{11} - R_{22}), \qquad 2P_{12} \sim \frac{1}{\delta_1^2}R_{12}$$

and

$$\lim_{\delta_1 \to \infty} [A_1^T, P] = \lim_{\delta_1 \to \infty}\begin{bmatrix} -2P_{12} & P_{11} - P_{22} \\ P_{11} - P_{22} & 2P_{12} \end{bmatrix} = 0,$$

which agrees with Lemma 3.9. Hence, $P_{11} - P_{22}$ and P_{12} both approach zero as $\delta_1 \to \infty$. These properties are the so-called equipartition (modal energy equilibration) and incoherence (modal decorrelation) phenomena [17, 20]. Since

$$\bar{\mu} \triangleq \lim_{\delta_1 \to \infty} \mu = \frac{1}{2\eta}\sqrt{\left(\frac{R_{11} - R_{22}}{2}\right)^2 + R_{12}^2},$$

the performance bound given by (43) approaches a (*finite*) constant as $\delta_1 \to \infty$. Furthermore, since $\lim_{\delta_1 \to \infty} P_{11} = \lim_{\delta_1 \to \infty} P_{22} = (1/4\eta)\,\mathrm{tr}\,(R)$, it follows that

$$\lim_{\delta_1 \to \infty} \mathrm{tr}\,(PV) + \mu\,\mathrm{tr}\,(V) = \left(\frac{1}{4\eta}\mathrm{tr}\,(R) + \bar{\mu}\right)\mathrm{tr}\,(V).$$

We now compare the performance bounds given by (39) and (43) for large values of δ_1. Denoting $\mathcal{T}_1 = \mathrm{tr}\,(V)\,\mathrm{tr}\,(P)$ and $\mathcal{T}_2 = \mathrm{tr}\,(PV) + \mu\,\mathrm{tr}\,(V)$, it can be shown using $R_{12}^2 < R_{11}R_{22}$ that

$$\lim_{\delta_1 \to \infty} \mathcal{T}_1 - \mathcal{T}_2 = \frac{\mathrm{tr}\,(V)}{2\eta}\left[\frac{R_{11} + R_{22}}{2} - \sqrt{\left(\frac{R_{11} - R_{22}}{2}\right)^2 + R_{12}^2}\right] = \frac{\mathrm{tr}\,(V)}{2\eta}\lambda_{\min}(R) > 0. \qquad (44)$$

Finally, if $\det R = 0$, then $\lim_{\delta_1 \to \infty} \mathcal{T}_1 = \lim_{\delta_1 \to \infty} \mathcal{T}_2 = (1/2\eta)\,\mathrm{tr}\,(V)\,\mathrm{tr}\,(R)$.

5. Numerical examples

Example 5.1. Let us consider a lightly damped system with $\zeta = 0.02$, $\omega_n = 2$, $\eta = \zeta\omega_n$, $\omega = \sqrt{1 - \zeta^2}\omega_n$,

$$A = \begin{bmatrix} -\eta & \omega \\ -\omega & -\eta \end{bmatrix}, \quad J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix},$$

and let

$$R = \begin{bmatrix} 2\beta & 0 \\ 0 & 2 \end{bmatrix},$$

where $\beta > 0$. For robust stability, we compare our result to the approach of [22]. For $R \neq 2I_2$ we must use a congruence transformation in order to apply the theorem in [22]. Hence, we transform

$$A^{\mathrm{T}}P + PA + R = 0 \tag{45}$$

to obtain

$$\hat{A}^{\mathrm{T}}\hat{P} + \hat{P}\hat{A} + 2I_2 = 0,$$

where $\hat{A} \triangleq S^{-1}AS$, and S is the congruence transformation matrix such that $S^{\mathrm{T}}RS = 2I_2$. As was mentioned in Remark 4.3, this system is robustly stable for all $\sigma_1 \in \mathbb{R}$. This follows from [22] by taking $\beta = 1$, that is, $R = 2I_2$, so that equation (45) has the solution $P = (1/\eta)I_2$. Therefore, in the notation of [22], $P_1 \triangleq \frac{1}{2}(J^{\mathrm{T}}P + PJ) = 0$, and thus the singular values of P_1 are all zero. As a result, the robust stability region is $|\sigma_1| < \infty$.

Now consider the case $\beta \gg 0$. Following the same procedure mentioned above, we have $|\sigma_1| \leq \delta_1 \sim (2/\omega\beta)(\eta^2 + \omega^2)$ as $\beta \to \infty$. Thus, for large β the approach of [22] becomes highly conservative. The reason for this conservatism is the similarity transformation of the skew-symmetric matrix J which was effectively imposed by the choice $R \neq 2I_2$. In the new basis, the matrix J is transformed to $S^{-1}JS$, which is no longer skew-symmetric.

Example 5.2. Consider the same system in Example 5.1 except with

$$R = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix},$$

and for robust performance, let

$$V = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

First, the robust stability region found by using the same technique as in the previous example is $|\sigma_1| < 1.37$, an extremely conservative result. As in the previous example, the reason for this conservatism is due to the similarity transformation of the skew-symmetric matrix J. In the new basis, the matrix J is transformed to $S^{-1}JS$, which is no longer skew-symmetric.

Next, let us compare the robust performance bound given by equation (39) in Proposition 4.2 with the bound suggested by Bernstein and Haddad [3]. According to (39) the performance bound is $\mathcal{T}(\mathcal{U}) \leq (1/2\eta)\operatorname{tr}(R) = 98.50$, which is valid for all $\sigma_1 \in \mathbb{R}$. In [3] the stability region and performance bound can be found by solving

$$A^{\mathrm{T}}P_A + P_A A + \Lambda + R = 0 \tag{46}$$

and by determining the values of σ_1 such that

$$\sigma_1(A_1^{\mathrm{T}}P_A + P_A A_1) \leq \Lambda, \tag{47}$$

where Λ is a nonnegative-definite matrix. First, letting $\Lambda = kI_2$, where $k \geq 0$, it can be shown that the solution to equation (46) is $P_A = P + (k/2\eta)I_2$, where P is the solution to (45) with

$$R = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.$$

Therefore, we have the performance bound $\mathcal{T}(\mathcal{U}) \leq \operatorname{tr}(PV) + (k/2\eta)\operatorname{tr}(V)$ with robust stability region $|\sigma_1| \leq k/\lambda_{\max}(J^{\mathrm{T}}P + PJ)$ (see Fig. 1). Alternatively, choosing $\Lambda = 0.53R$ yields the robust stability region $-2.57 \leq \sigma_1 \leq 0.37$ which yields the symmetric stability region $|\sigma_1| \leq 0.37$. For this robust stability region the performance bound $\mathcal{T}(U) \leq 118.20$ (see Fig. 2).
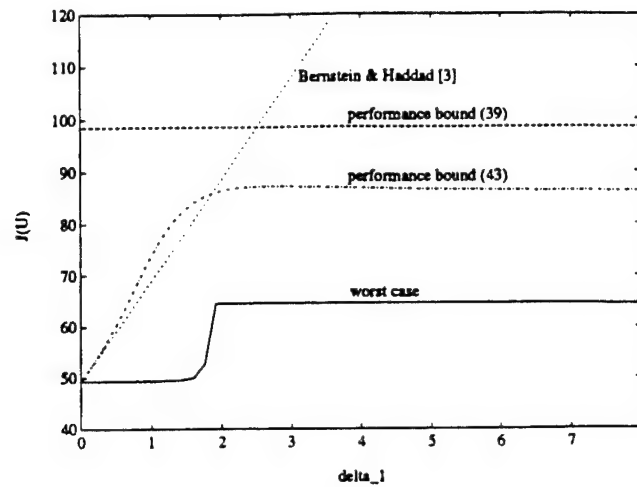
Fig. 1. Comparison of different robust performance bounds
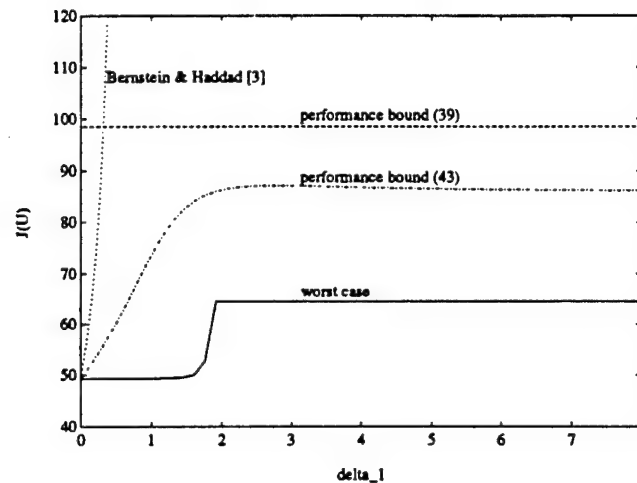
Fig. 2. Comparison of different robust performance bounds

6. Discussion and conclusions

As was shown in Propositions 4.2 and 4.6, the maximum-entropy-type Lyapunov functions correctly predict unconditional robust stability for arbitrary coordinates and thus, effectively, for an arbitrary state space basis. In addition, the performance bounds predicted by the maximum-entropy Lyapunov function are comparatively tight, even for large δ_1, whereas the bound of [3] is extremely conservative and highly coordinate-dependent. The problem of choosing an appropriate basis may be relatively benign if robust stability analysis is performed independently of robust performance analysis. That is, for robust stability analysis one can arbitrarily choose the state space basis to produce the best estimate of the robust stability region without regard to robust performance. However, in the problem of robust controller *synthesis* the basis is not arbitrary but rather is dictated by the weighting matrices V and R. Thus, the fact that the maximum-entropy-type Lyapunov functions provide robust stability and performance bounds that are only slightly affected by the choice of V and R appears to be a desirable feature for robust controller synthesis. This may explain the favorable results obtained in [2, 5, 6, 18, 19].

Acknowledgment

We wish to thank Jonathan How for noting Remark 4.5.

References

[1] D.S. Bernstein, Robust static and dynamic output-feedback stabilization: deterministic and stochastic perspectives, *IEEE Trans. Automat. Control* **32** (1987) 1076–1084.

[2] D.S. Bernstein and S.W. Greeley, Robust controller synthesis using the maximum entropy design equations, *IEEE Trans. Automat. Control* **13** (1986) 362–364.

[3] D.S. Bernstein and W.M. Haddad, Robust stability and performance analysis for linear dynamic systems, *IEEE Trans. Automat. Control* **34** (1989) 751–758.

[4] D.S. Bernstein and W.M. Haddad, Robust stability and performance analysis for state space system via Quadratic Lyapunov bounds, *SIAM J. Matrix Anal. Appl.* **11** (1990) 239–271.

[5] D.S. Bernstein and D.C. Hyland, The optimal projection/maximum entropy approach to designing low-order, robust controllers for flexible structures, in: *Proc. IEEE Conf. Dec. Contr.*, Fort Lauderdale, FL (1985) 745–752.

[6] D.S. Bernstein and D.C. Hyland, The optimal projection approach to robust, fixed-structure control design, in: J.L. Junkins, ed., *Mechanics and Control of Space Structures* (AIAA, New York, 1990) 287–293.

[7] J.W. Brewer, Kronecker products and matrix calculus in system, *IEEE Trans. Circuits and Systems* **25** (1978) 772–781.

[8] S.L. Campbell and C.D. Meyer Jr., *Generalized Inverse of Linear Transformation* (Pitman, New York, 1979).

[9] M. Cheung and S. Yurkovich, On the robustness of MEOP design versus asymptotic LQG synthesis, *IEEE Trans. Automat. Control* **33** (1988) 1061–1065.

[10] E.G. Collins Jr., J.A. King and D.S. Bernstein, Robust control design for the benchmark problem using the maximum entropy approach, in: *Proc. Amer. Contr. Conf.*, Boston, MA (1991) 1935–1936.

[11] E.G. Collins Jr., et al., High performance accelerometer-based control of the mini-MAST structure at Langley Research Center, NASA Contractor Report 4377, 1991.

[12] A. Gruzen, Robust reduced order control of flexible structures, C.S. Draper Laboratory Report CSDL-T-900, 1986.

[13] A. Gruzen and W.E. van der Velde, Robust reduced order control of flexible structures using the optimal projection/maximum entropy design methodology, in: *AIAA Guidance, Navigation, and Control Conf.*, Williamsburg, VA (1988).

[14] W.M. Haddad and D.S. Bernstein, Robust stabilization with positive real uncertainty: beyond the small gain theorem, *Systems Control Lett.* **17** (1991) 191–208.

[15] W.M. Haddad and D.S. Bernstein, Parameter-dependent Lyapunov functions, constant real parameter uncertainty, and the Popov criterion in robust analysis and synthesis, in: *Proc. IEEE Conf. Dec. Contr.*, Brighton (1991) 2274–2279 (Part I), 2632–2633 (Part II).

[16] N.W. Hagood IV and E.F. Crawley, Cost averaging techniques for robust control of parametrically uncertain system, MIT SERC Report #9-91, 1991.

[17] S.R. Hall, D.G. MacMartin and D.S. Bernstein, Covariance averaging in the analysis of uncertain systems, *IEEE Trans. Automat. Control*, to appear.

[18] D.C. Hyland, Maximum entropy stochastic approach to controller design for uncertain structural systems, in *Proc. American Control Conf.*, Arlington, VA (1982) 680–688.

[19] D.C. Hyland and A.N. Madiwale, A stochastic design approach for full-order compensation of structural systems with uncertain parameters, in: *Proc. AIAA Guidance and Control Conf.*, Albuquerque, NM (1981) 324–332.

[20] R.H. Lyon, *Statistical Energy Analysis of Dynamical Systems: Theory and Applications* (MIT Press, Cambridge, MA, 1975).

[21] W.M. Wonham, *Linear Multivariable Control: A Geometric Approach* (Springer, New York, 1974).

[22] K. Zhou and P.P. Khargonekar, Stability robustness bounds for linear state-space models with structured uncertainty, *IEEE Trans. Automat. Control* **32** (1987) 621–623.

Appendix B:

Homotopy Algorithm for Maximum Entropy Design

Homotopy Algorithm for Maximum Entropy Design

Emmanuel G. Collins Jr.,* Lawrence D. Davis,* and Stephen Richter†
Harris Corporation, Melbourne, Florida 32902

Maximum entropy design is a generalization of the LQG method that was developed to enable the synthesis of robust control laws for flexible structures. The method was developed by Hyland and motivated by insights gained from statistical energy analysis. Maximum entropy design has been used successfully in control design for ground-based structural testbeds and certain benchmark problems. The maximum entropy design equations consist of two Riccati equations coupled to two Lyapunov equations. When the uncertainty is zero, the equations decouple and the Riccati equations become the standard LQG regulator and estimator equations. A previous homotopy algorithm to solve the coupled equations relies on an iterative scheme that exhibits slow convergence properties as the uncertainty level is increased. This paper develops a new homotopy algorithm that does not suffer from this defect and in fact can have quadratic convergence rates along the homotopy curve. Algorithms of this type should also prove effective in the solution of other sets of coupled Riccati and Lyapunov equations appearing in robust control theory.

Nomenclature

$e_m^{(i)}$	= m-dimensional column vector whose ith element equals one and whose additional elements are zeros
I_r	= $r \times r$ identity matrix
$\mathfrak{R}^n$, $\mathfrak{R}^{m \times n}$	= $n \times 1$ real vectors, $m \times n$ real matrices
tr Z	= trace of square matrix Z
vec($\cdot$)	= invertible linear operator defined such that vec$(S) \triangleq [s_1^T s_2^T \cdots s_q^T]^T$, $S \in \mathfrak{R}^{p \times q}$ where $s_j \in \mathfrak{R}^p$ denotes the jth column of S
$Y > Z$	= $Y - Z$ is positive definite
$Y \geq Z$	= $Y - Z$ is nonnegative definite
Y/Z	= matrix whose (i, j) element is y_{ij}/z_{ij}, Y and Z must have identical dimensions (MATLAB notation)
$Y*Z$	= Hadamard product of Y and Z ($[y_{ij}z_{ij}]$), Y and Z must have identical dimensions
$Z*$	= complex conjugate of the matrix Z
Z^H	= complex conjugate transpose of the matrix Z, $(Z*)^T$
$Z(k, :)$	= kth row of the matrix Z (MATLAB notation)
$Z(:, k)$	= kth column of the matrix Z (MATLAB notation)
z_{ij}, $Z_{i,j}$, or $Z_{(i,j)}$	= (i, j) element of matrix Z
$\otimes$	= Kronecker product[14]

I. Introduction

THE linear-quadratic-Gaussian (LQG) compensator[1] has been developed to facilitate the design of control laws for complex, multi-input/multi-output (MIMO) systems such as flexible structures. However, it is well known that an LQG compensator can yield a closed-loop system with arbitrarily poor robustness properties.[2] This deficiency has led to generalizations of LQG that allow the design of robust controllers. One such generalization of LQG is the maximum entropy control design approach that was originated by Hyland[3] and Bern-

stein and Hyland.[4,5] Maximum entropy control design was developed specifically to enable robust control law design for flexible structures. In particular, this design technique develops control laws that are insensitive to changes in the (undamped) modal frequencies. The approach was motivated by insights from statistical energy analysis and has proven to be an effective tool in the design of robust control laws for ground-based flexible structure testbeds[6,7] and for certain benchmark problems.[8-10]

The rigorous theoretical foundation for maximum entropy design is not yet complete. However, in Ref. 11 it is shown that, for an open-loop system, a Lyapunov function based on the maximum entropy constraint equation predicts unconditional stability for changes in the undamped natural frequency. The results of Ref. 11 also provide evidence that the theoretical foundation of maximum entropy analysis and design may be related to recent robustness results based on parameter-dependent Lyapunov functions.[12]

The computation of full-order maximum entropy controllers requires the solution of a set of equations consisting of two Riccati equations coupled to two Lyapunov equations. If the uncertainty is assumed to be zero, these equations decouple and the Riccati equations become the standard LQG Riccati equations. A homotopy algorithm for solving these equations is described in Ref. 13. This algorithm is based on first solving an LQG problem and gradually increasing the uncertainty level until the desired degree of robustness is achieved. Unfortunately, the algorithm of Ref. 13 relies on an iterative scheme that tends to have increasingly poor convergence properties as the uncertainty level is increased.

The contribution of this paper is the development of a new homotopy algorithm for full-order maximum entropy design. Unlike the previous approach, this algorithm can have quadratic convergence rates along the homotopy curve. Algorithms of this type should also prove effective in the solution of other sets of coupled Riccati and Lyapunov equations appearing in robust control theory (e.g., Ref. 12). The algorithm has been implemented in MATLAB and is illustrated using a control problem from the Active Control Technique Evaluation for Spacecraft (ACES) testbed at NASA Marshall Space Flight Center in Huntsville, Alabama. A useful feature of maximum entropy design, seen in the example, is that it often produces controllers that are effectively reduced-order controllers. Other features of maximum entropy controllers are described in Refs. 6 and 7.

The paper is organized as follows. Section II develops the maximum entropy design equations. Section III gives a brief synopsis of homotopy methods. Next, Sec. IV develops a new

homotopy algorithm for maximum entropy control design. Section V illustrates the algorithm using a 17th-order model of one of the transfer functions of the ACES structure at NASA Marshall Space Flight Center. Finally, Sec. VI discusses the conclusions.

II.   Maximum Entropy Design Equations
Consider the system

$$\dot{x}(t) = Ax(t) + Bu(t) + w_1(t)$$

$$y(t) = Cx(t) + Du(t) + w_2(t)$$

where $x \in \mathbb{R}^{n_x}$, $u \in \mathbb{R}^{n_u}$, $y \in \mathbb{R}^{n_y}$, $w_1 \in \mathbb{R}^{n_x}$ is white disturbance noise with intensity $V_1 \geq 0$, $w_2 \in \mathbb{R}^{n_y}$ is white observation noise with intensity $V_2 > 0$, and w_1 and w_2 have cross correlation $V_{12} \in \mathbb{R}^{n_x \times n_y}$. It is assumed that (A, B) is stabilizable and (A, C) is detectable. Also, the matrix A is assumed to be of the form

$$A = \text{block diag}[A^{(1)}, A^{(2)}]$$

where $A^{(2)}$ represents the dynamics that are certain and $A^{(1)}$ represents the nominal dynamics of the uncertain modes and is in real normal form; for example,

$$A^{(1)} = \text{block diag}\left\{\begin{bmatrix} -\nu_1 & \omega_1 \\ -\omega_1 & -\nu_1 \end{bmatrix}, \quad -\nu_2, \quad \begin{bmatrix} -\nu_3 & \omega_3 \\ -\omega_3 & -\nu_3 \end{bmatrix}\right\}$$

We also assume that only the modes with complex eigenvalues, corresponding to the 2×2 blocks

$$\begin{bmatrix} -\nu_j & \omega_j \\ -\omega_j & -\nu_j \end{bmatrix}$$

are uncertain and that the uncertainty patterns $A_i \in \mathbb{R}^{n_x \times n_x}$ are of the form

$$A_i = \text{block diag}\left\{0, \ldots, 0, \quad \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad 0, \ldots, 0\right\}$$

Notice that the A_i correspond to errors in the undamped natural frequencies, i.e., the imaginary part of the eigenvalues.

The maximum entropy control design problem is stated as follows. Find a full-order dynamic compensator (i.e., a compensator of order n_x),

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t)$$

$$u(t) = -C_c x_c(t)$$

which stabilizes $\bar{A}_s$, defined later, and minimizes the cost functional

$$J(A_c, B_c, C_c) = \text{tr } \bar{Q}\bar{R}$$

where $\bar{Q}$ satisfies

$$0 = \bar{A}_s \bar{Q} + \bar{Q}\bar{A}_s^T + \bar{V} + \sum_{i=1}^{n_a} \bar{A}_i \bar{Q}\bar{A}_i^T$$

and

$$\bar{A}_s = \bar{A} + \frac{1}{2}\sum_{i=1}^{n_a} \alpha_i^2 \bar{A}_i^2, \qquad \bar{A}_i = \text{block diag}\{A_i, 0_{n_x}\}$$

$$\bar{A} = \begin{bmatrix} A & -BC_c \\ B_c C & A_c - B_c DC_c \end{bmatrix}$$

$$\bar{R} = \begin{bmatrix} R_1 & R_{12}C_c \\ C_c^T R_{12}^T & C_c^T R_2 C_c \end{bmatrix}, \qquad \bar{V} = \begin{bmatrix} V_1 & V_{12}B_c^T \\ B_c V_{12}^T & B_c V_2 B_c^T \end{bmatrix}$$

There is currently no rigorous justification for the requirement that $\bar{A}_s$ be stabilized, but extensive numerical examples have shown that stability of $\bar{A}_s$ insures stability of the nominal closed-loop system. Notice that if no uncertainty is assumed (i.e., $\alpha_i \triangleq 0$), then the maximum entropy control design problem becomes the standard LQG problem. The solution to the maximum entropy problem is characterized by the following theorem.

Theorem 1[3-5]. Suppose (A_c, B_c, C_c) solves the maximum entropy control design problem. Then, there exist nonnegative-definite matrices Q, P, $\hat{Q}$, and $\hat{P}$ such that A_c, B_c, and C_c are given by

$$A_c = A_s - BR_2^{-1}P_a - Q_a V_2^{-1}C + Q_a V_2^{-1}DR_2^{-1}P_a$$

$$B_c = Q_a V_2^{-1}, \qquad C_c = R_2^{-1}P_a$$

where

$$A_s = A + \frac{1}{2}\sum_{i=1}^{n_a} \alpha_i^2 A_i^2$$

$$P_a = B^T P + R_{12}^T, \qquad Q_a = QC^T + V_{12}$$

and the following conditions are satisfied:

$$0 = A_s^T P + PA_s + R_1 - P_a^T R_2^{-1}P_a + \sum_{i=1}^{n_a} \alpha_i^2 A_i^T(P + \hat{P})A_i \quad (1)$$

$$0 = A_s Q + QA_s^T + V_1 - Q_a V_2^{-1}Q_a^T + \sum_{i=1}^{n_a} \alpha_i^2 A_i(Q + \hat{Q})A_i^T \quad (2)$$

$$0 = (A_s - Q_a V_2^{-1}C)^T\hat{P} + \hat{P}(A_s - Q_a V_2^{-1}C) + P_a^T R_2^{-1}P_a \quad (3)$$

$$0 = (A_s - BR_2^{-1}P_a)\hat{Q} + \hat{Q}(A_s - BR_2^{-1}P_a)^T + Q_a V_2^{-1}Q_a^T \quad (4)$$

Remark 1. If no uncertainty is assumed (i.e., $\alpha_i \triangleq 0$), then Eqs. (1–4) decouple, Eqs. (1) and (2) become the standard LQG regulator and estimator Riccati equations, and (A_c, B_c, C_c) defined in Theorem 1 is an LQG compensator.

III.   Homotopy Methods for the Solution of Nonlinear Algebraic Equations
In the next section, we present a homotopy algorithm for solving the maximum entropy design equations (1–4). A homotopy is a continuous deformation of one function into another. The purpose of this section is to provide a very brief description of homotopy methods for finding the solutions of nonlinear algebraic equations. The reader is referred to Refs. 15–17 for additional details.

The basic problem is as follows. Given set Θ and Φ contained in $\mathbb{R}^n$ and a mapping $F : \Theta \to \Phi$, find solutions to

$$F(\theta) = 0$$

Homotopy methods embed the problem $F(\theta) = 0$ in a larger problem. In particular, let $H : \Theta \times [0, 1] \to \mathbb{R}^n$ be such that the following conditions exist:
1) $H(\theta, 1) = F(\theta)$.
2) There exists at least one known $\theta_0 \in \mathbb{R}^n$ that is a solution to $H(\cdot, 0) = 0$, i.e.,

$$H(\theta_0, 0) = 0$$

3) There exists a continuous curve $(\theta(\lambda), \lambda)$ in $\mathbb{R}^n \times [0, 1]$ such that

$$H(\theta(\lambda), \lambda) = 0 \quad \text{for} \quad \lambda \in [0, 1]$$

with

$$(\theta(0), 0) = (\theta_0, 0)$$

4) The curve $(\theta(\lambda), \lambda)$ is differentiable.

A homotopy algorithm then constructs a procedure to compute the actual curve $(\theta(\lambda), \lambda)$ such that the initial solution $\theta(0)$ is transformed to a desired solution $\theta(1)$ satisfying

$$0 = H(\theta(1), 1) = F(\theta(1))$$

Differentiating $H(\theta(\lambda), \lambda) = 0$ with respect to λ yields Davidenko's differential equation:

$$\frac{\partial H}{\partial \theta}\frac{d\theta}{d\lambda} + \frac{\partial H}{\partial \lambda} = 0 \tag{5}$$

Together with $\theta(0) = \theta_0$, Eq. (5) defines an initial value problem that by numerical integration from 0 to 1 yields the desired solution $\theta(1)$. Some numerical integration schemes are described in Ref. 17.

IV. Homotopy Algorithm for Full-Order Maximum Entropy Control Design

This section presents a novel homotopy algorithm that can be used to design full-order maximum entropy controllers. The algorithm is based on explicitly solving the four coupled maximum entropy design equations given in Eqs. (1–4).

A. Homotopy Map

To define the homotopy map we assume that the plant matrices (A, B, C, D), the cost-weighting matrices (R_1, R_2, R_{12}), the disturbance matrices (V_1, V_2, V_{12}), and the vector of uncertainty weights $(\alpha \in \mathcal{R}^{n_\alpha})$ are functions of the homotopy parameter $\lambda \in [0, 1]$. In particular, the following is assumed:

$$\begin{bmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} = \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} + \lambda \left\{ \begin{bmatrix} A_f & B_f \\ C_f & D_f \end{bmatrix} - \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} \right\}$$

$$\begin{bmatrix} R_1(\lambda) & R_{12}(\lambda) \\ R_{12}^T(\lambda) & R_2(\lambda) \end{bmatrix} = L_R(\lambda) L_R^T(\lambda)$$

where

$$L_R(\lambda) = L_{R,0} + \lambda(L_{R,f} - L_{R,0})$$

and $L_{R,0}$ and $L_{R,f}$ satisfy

$$L_{R,0}L_{R,0}^T \triangleq \begin{bmatrix} R_{1,0} & R_{12,0} \\ R_{12,0}^T & R_{2,0} \end{bmatrix}, \quad L_{R,f}L_{R,f}^T \triangleq \begin{bmatrix} R_{1,f} & R_{12,f} \\ R_{12,f}^T & R_{2,f} \end{bmatrix}$$

$$\begin{bmatrix} V_1(\lambda) & V_{12}(\lambda) \\ V_{12}^T(\lambda) & V_2^T(\lambda) \end{bmatrix} = L_V(\lambda) L_V^T(\lambda)$$

where

$$L_V(\lambda) = L_{V,0} + \lambda(L_{V,f} - L_{V,0})$$

and $L_{V,0}$ and $L_{V,f}$ satisfy

$$L_{V,0}L_{V,0}^T = \begin{bmatrix} V_{1,0} & V_{12,0} \\ V_{12,0}^T & V_{2,0} \end{bmatrix}, \quad L_{V,f}L_{V,f}^T = \begin{bmatrix} V_{1,f} & V_{12,f} \\ V_{12,f}^T & V_{2,f} \end{bmatrix}$$

$$\alpha_i^2(\lambda) = \alpha_{0,i}^2 + \lambda(\alpha_{f,i}^2 - \alpha_{0,i}^2), \qquad i = 1, 2, \ldots, n_\alpha$$

Notice that at $\lambda = 0$, $A(\lambda) = A_0$, $B(\lambda) = B_0, \ldots, \alpha_i^2(\lambda) = \alpha_{0,i}^2$, whereas at $\lambda = 1$, $A(\lambda) = A_f$, $B(\lambda) = B_f, \ldots, \alpha_i^2(\lambda) = \alpha_{f,i}^2$. Some guidelines for choosing the initial and final matrices are discussed later in Sec. IV.C.

The homotopy $0 = H((P, Q, \hat{P}, \hat{Q}), \lambda)$ is given by the equations

$$0 = A_s(\lambda)^T P(\lambda) + P(\lambda)A_s(\lambda) + R_1(\lambda) - P_a(\lambda)^T R_2(\lambda)^{-1} P_a(\lambda)$$

$$+ \sum_{i=1}^{n_\alpha} \alpha_i^2(\lambda) A_i^T P(\lambda) A_i + \sum_{i=1}^{n_\alpha} \alpha_i^2(\lambda) A_i^T \hat{P}(\lambda) A_i \tag{6}$$

$$0 = A_s(\lambda)Q(\lambda) + Q(\lambda)A_s(\lambda)^T + V_1(\lambda) - Q_a(\lambda)V_2^{-1}(\lambda)Q_a(\lambda)^T$$

$$+ \sum_{i=1}^{n_\alpha} \alpha_i^2(\lambda)A_i Q(\lambda)A_i^T + \sum_{i=1}^{n_\alpha} \alpha_i^2 A_i \hat{Q}(\lambda)A_i^T \tag{7}$$

$$0 = \left[A_s(\lambda) - Q_a(\lambda)V_2^{-1}(\lambda)C(\lambda)\right]^T \hat{P}(\lambda)$$

$$+ \hat{P}(\lambda)\left[A_s(\lambda) - Q_a(\lambda)V_2^{-1}(\lambda)C(\lambda)\right]$$

$$+ P_a(\lambda)^T R_2^{-1}(\lambda)P_a(\lambda) \tag{8}$$

$$0 = \left[A_s(\lambda) - B(\lambda)R_2^{-1}(\lambda)P_a(\lambda)\right]\hat{Q}(\lambda)$$

$$+ \hat{Q}(\lambda)\left[A_s(\lambda) - B(\lambda)R_2^{-1}(\lambda)P_a(\lambda)\right]^T$$

$$+ Q_a(\lambda)V_2^{-1}(\lambda)Q_a(\lambda)^T \tag{9}$$

where

$$A_s(\lambda) \triangleq A(\lambda) + \frac{1}{2}\sum_{i=1}^{n_\alpha} \alpha_i^2(\lambda)A_i^2$$

$$P_a(\lambda) \triangleq B(\lambda)^T P(\lambda) + R_{12}(\lambda)^T, \quad Q_a(\lambda) \triangleq Q(\lambda)C(\lambda)^T + V_{12}(\lambda)$$

B. Derivative and Correction Equations

The homotopy algorithm presented in the next section uses a predictor/corrector numerical integration scheme. The predictor steps require derivatives $[\dot{P}(\lambda), \dot{Q}(\lambda), \dot{\hat{P}}(\lambda) \dot{\hat{Q}}(\lambda)]$, where $\dot{M} \triangleq dM/d\lambda$, whereas the correction step is based on using Newton corrections, denoted here as $(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q})$. Next we derive the matrix equations that can be used to solve for the derivatives and corrections. For notational simplicity we omit the argument λ in the derived equations.

1. Derivative Equations

Differentiating Eqs. (6–9) with respect to λ gives the following coupled matrix equations:

$$0 = A_P^T \dot{P} + \dot{P}A_P + R + \sum_{i=1}^{n_\alpha} \alpha_i^2 A_i^T \dot{P}A_i + \sum_{i=1}^{n_\alpha} \alpha_i^2 A_i^T \dot{\hat{P}}A_i \tag{10}$$

$$0 = A_Q\dot{Q} + \dot{Q}A_Q^T + V + \sum_{i=1}^{n_\alpha} \alpha_i^2 A_i \dot{Q}A_i^T + \sum_{i=1}^{n_\alpha} \alpha_i^2 A_i \dot{\hat{Q}}A_i^T \tag{11}$$

$$0 = A_Q^T \dot{\hat{P}} + \dot{\hat{P}}A_Q + \hat{R} + G_C\dot{Q}\hat{F} + \hat{F}\dot{Q}G_C$$

$$+ H_P^T \dot{P}K_P + K_P^T \dot{P}H_P \tag{12}$$

$$0 = A_P\dot{\hat{Q}} + \dot{\hat{Q}}A_P^T + \hat{V} + G_B\dot{P}\hat{E} + \hat{E}\dot{P}G_B$$

$$+ H_Q\dot{Q}K_Q^T + K_Q\dot{Q}H_Q^T \tag{13}$$

where

$$A_P \triangleq A_s - BR_{2,\text{inv}}P_a, \qquad A_Q \triangleq A_s - Q_a V_{2,\text{inv}}C$$

$$R \triangleq \dot{A}_s^T P + P\dot{A}_s + \dot{R}_1 - P_a^T R_{2,\text{inv}}(\dot{B}^T P + \dot{R}_{12}^T)$$

$$- (P\dot{B} + \dot{R}_{12})R_{2,\text{inv}}P_a - P_a^T \dot{R}_{2,\text{inv}}P_a$$

$$+ \sum_{i=1}^{n_\alpha} \dot{\alpha}_{i,\text{sq}}A_i^T (P + \hat{P})A_i$$

$$V \triangleq \dot{A}_s Q + Q\dot{A}_s^T + \dot{V}_1 - Q_a V_{2,\text{inv}}(\dot{C}Q + \dot{V}_{12}^T)$$

$$- (Q\dot{C}^T + \dot{V}_{12})V_{2,\text{inv}}Q_a^T - Q_a \dot{V}_{2,\text{inv}}Q_a^T$$

$$+ \sum_{i=1}^{n_\alpha} \dot{\alpha}_{i,\text{sq}}A_i(Q + \hat{Q})A_i^T$$

$$\hat{R} \triangleq \left[\dot{A}_s - Q_a V_{2,\text{inv}}\dot{C} - Q_a \dot{V}_{2,\text{inv}}C - (Q\dot{C}^T + \dot{V}_{12})V_{2,\text{inv}}C\right]^T \hat{P}$$

$$+ \hat{P}\left[\dot{A}_s - Q_a V_{2,\text{inv}}\dot{C} - Q_a \dot{V}_{2,\text{inv}}C - (Q\dot{C}^T + \dot{V}_{12})V_{2,\text{inv}}C\right]$$

$$+ P_a^T R_{2,\text{inv}}(\dot{B}^T P + \dot{R}_{12}^T) + (\dot{B}^T P + \dot{R}_{12}^T)^T R_{2,\text{inv}} P_a$$

$$+ P_a^T \dot{R}_{2,\text{inv}} P_a$$

$$\hat{V} \triangleq \left[\dot{A}_s - \dot{B}R_{2,\text{inv}}P_a - B\dot{R}_{2,\text{inv}}P_a - BR_{2,\text{inv}}(\dot{B}^T P + \dot{R}_{12}^T)\right]\hat{Q}$$

$$+ \hat{Q}\left[\dot{A}_s - \dot{B}R_{2,\text{inv}}P_a - B\dot{R}_{2,\text{inv}}P_a - BR_{2,\text{inv}}(\dot{B}^T P + \dot{R}_{12}^T)\right]^T$$

$$+ Q_a V_{2,\text{inv}}(Q\dot{C}^T + \dot{V}_{12})^T + (Q\dot{C}^T + \dot{V}_{12})V_{2,\text{inv}}Q_a^T$$

$$+ Q_a \dot{V}_{2,\text{inv}} Q_a^T$$

$$G_B = -BR_{2,\text{inv}}B^T, \quad G_c = -C^T V_{2,\text{inv}}C, \quad \hat{E} = \hat{Q}, \quad \hat{F} = \hat{P}$$

$$H_P = BR_{2,\text{inv}}P_a, \quad H_Q = Q_a V_{2,\text{inv}}C, \quad K_P = I_{n_x}, \quad K_Q = I_{n_x}$$

Note that in the previous equations we have used the notations

$$R_{2,\text{inv}} \triangleq R_2^{-1}, \qquad V_{2,\text{inv}} \triangleq V_2^{-1}, \qquad \alpha_{i,\text{sq}} \triangleq \alpha_i^2$$

2. Correction Equations

The correction equations are developed with λ at some fixed value, say λ^*. The derivation of the correction equations is based on the relationship between Newton's method and a particular homotopy. In the following text we use the notation

$$f'(\theta) \triangleq \frac{\partial f}{\partial \theta}$$

Let $f : \Re^n \to \Re^n$ be C^1 and consider the equation

$$0 = f(\theta) \tag{14}$$

If $\theta^{(i)}$ is the current approximation to the solution of Eq. (14), then the Newton correction[18] $\Delta\theta$ is given by

$$\theta^{(i+1)} - \theta^{(i)} \triangleq \Delta\theta = -f'(\theta^{(i)})^{-1}e \tag{15}$$

where

$$e \triangleq f(\theta^{(i)})$$

Now, let $\theta^{(i)}$ be an approximation to θ satisfying Eq. (14). Then, with e as given immediately above, construct the following homotopy to solve Eq. (14):

$$(1-\beta)e = f(\theta(\beta)), \qquad \beta \in [0, 1] \tag{16}$$

[Note that at $\beta = 0$ Eq. (16) has solution $\theta(0) = \theta^{(i)}$, whereas $\theta(1)$ satisfies Eq. (14)]. Then, differentiating Eq. (16) with respect to β gives

$$\left.\frac{\partial\theta}{\partial\beta}\right|_{\beta=0} = -f'(\theta^{(i)})^{-1}e \tag{17}$$

Remark 2. Note that the Newton correction $\Delta\theta$ in Eq. (15) and the derivative $\partial\theta/\partial\beta|_{\beta=0}$ in Eq. (17) are identical. Hence, the Newton correction $\Delta\theta$ can be found by constructing a homotopy of the form of Eq. (16) and solving for the resulting derivative $\partial\theta/\partial\beta|_{\beta=0}$. As seen later, this insight is particularly useful when deriving Newton corrections for equations that have a matrix structure. It is also of interest to note that the homotopy of Eq. (16) is appropriately referred to in some literature as a "Newton homotopy."[15]

Now, we use the insights of Remark 2 to derive the equations that need to be solved for the Newton corrections $(\Delta P, \Delta Q, \Delta\hat{P}, \Delta\hat{Q})$. We begin by recalling that λ is assumed to have some fixed value, say λ^*. Also, it is assumed that P^*, Q^*,

$\hat{P}^*$, and $\hat{Q}^*$ are the current approximations to $P(\lambda^*)$, $Q(\lambda^*)$, $\hat{P}(\lambda^*)$, and $\hat{Q}(\lambda^*)$ and that E_P, E_Q, $E_{\hat{P}}$, and $E_{\hat{Q}}$ are, respectively, the errors in Eqs. (1–4) with $\lambda = \lambda^*$ and $P(\lambda)$, $Q(\lambda)$, $\hat{P}(\lambda)$, and $\hat{Q}(\lambda)$ replaced by P^*, Q^*, $\hat{P}^*$, and $\hat{Q}^*$.

We next form the homotopy

$$(1-\beta)E_P = A_s^T P(\beta) + P(\beta)A_s + R_1 - P_a(\beta)^T R_2^{-1} P_a(\beta)$$

$$+ \sum_{i=1}^{n_a} \alpha_i^2 A_i^T P(\beta)A_i + \sum_{i=1}^{n_a} \alpha_i^2 A_i^T \hat{P}(\beta)A_i \tag{18}$$

$$(1-\beta)E_Q = A_s Q(\beta) + Q(\beta)A_s^T + V_1 - Q_a(\beta)V_2^{-1}Q_a(\beta)^T$$

$$+ \sum_{i=1}^{n_a} \alpha_i^2 A_i Q(\beta)A_i^T + \sum_{i=1}^{n_a} \alpha_i^2 A_i \hat{Q}(\beta)A_i^T \tag{19}$$

$$(1-\beta)E_{\hat{P}} = \left[A_s Q_a(\beta)V_2^{-1}C\right]^T \hat{P}(\beta)$$

$$+ \hat{P}(\beta)\left[A_s - Q_a(\beta)V_2^{-1}C\right] + P_a(\beta)^T R_2^{-1} P_a(\beta) \tag{20}$$

$$(1-\beta)E_{\hat{Q}} = \left[A_s - BR_2^{-1}P_a(\beta)\right]\hat{Q}(\beta)$$

$$+ \hat{Q}(\beta)\left[A_s - BR_2^{-1}P_a(\beta)\right]^T + Q_a(\beta)V_2^{-1}Q_a(\beta)^T \tag{21}$$

where

$$A_s = A + \sum_{i=1}^{n_a} \alpha_i^2 A_i^2$$

$$P_a = B^T P(\beta) + R_{12}^T, \qquad Q_a = Q(\beta)C^T + V_{12}$$

and the system matrices are assumed to be evaluated at $\lambda = \lambda^*$, i.e., $(A, B, \ldots, R_1, R_2, \ldots) = [A(\lambda^*), B(\lambda^*), \ldots, R_1(\lambda^*), R_2(\lambda^*), \ldots]$. Differentiating Eqs. (18–21) with respect to β and using Remark 4 to make the replacements

$$\Delta P = \left.\frac{dP}{d\beta}\right|_{\beta=0}, \qquad \Delta Q = \left.\frac{dQ}{d\beta}\right|_{\beta=0}$$

$$\Delta\hat{P} = \left.\frac{d\hat{P}}{d\beta}\right|_{\beta=0}, \qquad \Delta\hat{Q} = \left.\frac{d\hat{Q}}{d\beta}\right|_{\beta=0}$$

gives

$$0 = A_P^T \Delta P + \Delta P A_P + R + \sum_{i=1}^{n_a} \alpha_i^2 A_i^T \Delta P A_i$$

$$+ \sum_{i=1}^{n_a} \alpha_i^2 A_i^T \Delta\hat{P}A_i \tag{22}$$

$$0 = A_Q \Delta Q + \Delta Q A_Q^T + V + \sum_{i=1}^{n_a} \alpha_i^2 A_i \Delta Q A_i^T$$

$$+ \sum_{i=1}^{n_a} \alpha_i^2 A_i \Delta\hat{Q}A_i^T \tag{23}$$

$$0 = A_Q^T \Delta\hat{P} + \Delta\hat{P}A_Q + \hat{R} + G_C \Delta Q\hat{F} + \hat{F}\Delta Q G_C$$

$$+ H_P^T \Delta P K_P + K_P^T \Delta P H_P \tag{24}$$

$$0 = A_P \Delta\hat{Q} + \Delta\hat{Q}A_P^T + \hat{V} + G_B \Delta P\hat{E} + \hat{E}\Delta P G_B$$

$$+ H_Q \Delta Q K_Q^T + K_Q \Delta Q H_Q^T \tag{25}$$

where

$$A_P \triangleq A_s - BR_2^{-1}P_a, \qquad A_Q \triangleq A_s - Q_a V_2^{-1}C$$

$$R = E_P, \qquad V = E_Q, \qquad \hat{R} = E_{\hat{P}}, \qquad \hat{V} = E_{\hat{Q}}$$

$$G_B = -BR_2^{-1}B^T, \quad G_c = -C^T V_2^{-1}C, \quad \hat{E} = \hat{Q}, \quad \hat{F} = \hat{P}$$

$$H_P = BR_2^{-1}P_a, \quad H_Q = Q_a V_2^{-1}C, \quad K_P = I_{n_x}, \quad K_Q = I_{n_x}$$

Comparing Eqs. (22-25) with Eqs. (10-13) reveals that the derivative and correction equations are identical in form. Each set of equations consists of four coupled Lyapunov equations. Since these equations are linear, by using Kronecker products[14] they can be converted to the vector form $\mathcal{C}\chi = b$ where for Eqs. (22-25) χ is a vector containing the independent elements of ΔP, ΔQ, $\Delta \hat{P}$, and $\Delta \hat{Q}$. The $\mathcal{C}$ is then a square matrix of dimension $2n_x(n_x + 1)$. Inversion of $\mathcal{C}$ is hence very computationally intensive for even relatively small problems (e.g., $n_x = 10$).

Fortunately, the coupling terms described by the summation terms in Eqs. (22) and (23) are relatively sparse. In particular, each summation has only $3n_\alpha$ independent terms. Hence, a technique similar to that described in Ref. 19, which exploits this sparseness, can be used to efficiently solve Eqs. (22-25) [or equivalently Eqs. (10-13)]. The details of the solution procedure are described in Appendix B. The solution procedure relies on the solution of a maximum entropy Lyapunov equation as described in Appendix A. The results of Appendix A are also based on the results of Ref. 19. Both Appendices A and B rely on diagonalization of the coefficient matrices of each of the Lyapunov equations. Since efficient MATLAB implementation requires the minimization of the use of *for loops*, the solution procedures of Appendices A and B implement the techniques of Ref. 19 with minimal looping. A complete derivation of these results is presented in Ref. 20.

C.   Overview of the Homotopy Algorithm

This section describes the general logic and features of the homotopy algorithm for full-order maximum entropy control. It is assumed that the designer has supplied a set of system matrices $S_f = (A_f, B_f, C_f, D_f, R_{1,f}, R_{2,f}, R_{12,f}, V_{1,f}, V_{2,f}, V_{12,f}, \alpha_f)$ describing the optimization problem whose solution is desired. In addition, it is assumed that the designer has chosen an initial set of related system matrices $S_0 = (A_0, B_0, C_0, D_0, R_{1,0}, R_{2,0}, R_{12,0}, V_{1,0}, V_{2,0}, V_{1,0}, \alpha_0)$ that has an easily obtained or known solution $(P_0, Q_0 \hat{P}_0, \hat{Q}_0)$ to the maximum entropy design equations. Note that we can always choose $\alpha_0 = 0$ in which case $(P_0, Q_0, \hat{P}_0, \hat{Q}_0)$ corresponds to an LQG problem and can be computed using standard Riccati equation and Lyapunov equation solvers. In practice, we often choose the remaining system matrices to have equal initial and final values, i.e., $A_f = A_0$, $B_f = B_0$, ..., $R_{1,f} = R_{1,0}$, $R_{2,f} = R_{2,0}$, ..., $V_{1,f} = V_{1,0}$, $V_{2,f} = V_{2,0}$. However, there is a strong rationale for allowing these matrices to vary during the homotopy. For example, suppose a maximum entropy controller of a particular robustness (corresponding to some value of α) is designed but the controller authority level is not desirable. Then, instead of changing the weights R_1, R_2, R_{12}, V_1, V_2, and V_{12} to reflect the desired authority level, solving the corresponding LQG problem (that is, the problem with $\alpha = 0$), and then using the homotopy algorithm to reinsert the robustness (corresponding to the original value of α), we can use the homotopy algorithm to modify the weights R_1, R_2, ..., with α fixed to its original value. Similarly, we can modify the nominal plant matrices A, B, C, and D with α fixed to reflect new data concerning the plant.

Later we present an outline of the homotopy algorithm. This algorithm describes a predictor/corrector numerical integration scheme. The prediction step uses cubic spline prediction as described next.

1.   Cubic Spline Prediction

Here we use the notation that λ_0, λ_{-1}, and λ_1 represent the values of λ at, respectively, the current point on the homotopy curve, the previous point, and the next point. Also, $\dot{M} \triangleq dM/d\lambda$. The prediction of $P(\lambda_1)$ requires $P(\lambda_0)$, $\dot{P}(\lambda_0)$, $P(\lambda_{-1})$, and $\dot{P}(\lambda_{-1})$. In particular,

$$\text{vec}\left[P(\lambda_1)\right] = a_0 + a_1\lambda_1 + a_2\lambda_1^2 + a_3\lambda_1^3$$

where a_0, a_1, a_2, and a_3 are computed by solving

$$[a_0 \; a_1 \; a_2 \; a_3] \begin{bmatrix} 1 & 0 & 1 & 0 \\ \lambda_{-1} & 1 & \lambda_0 & 1 \\ \lambda_{-1}^2 & 2\lambda_{-1} & \lambda_0^2 & 2\lambda_0 \\ \lambda_{-1}^3 & 3\lambda_{-1}^2 & \lambda_0^3 & 3\lambda_0^2 \end{bmatrix} = \begin{Bmatrix} \text{vec}\left[P(\lambda_{-1})\right] \\ \text{vec}\left[\dot{P}(\lambda_{-1})\right] \\ \text{vec}\left[P(\lambda_0)\right] \\ \text{vec}\left[\dot{P}(\lambda_0)\right] \end{Bmatrix}$$

Note that if $P(\lambda_{-1})$ and $\dot{P}(\lambda_{-1})$ are not available (as occurs at the initial iteration of the homotopy algorithm), the $P(\lambda_1)$ is predicted using linear prediction, i.e.,

$$P(\lambda_1) = P(\lambda_0) + (\lambda_1 - \lambda_0)\dot{P}(\lambda_0)$$

2.   Outline of the Homotopy Algorithm

Step 1:   Initialize loop = 0, $\lambda = 0$, $\Delta\lambda \in [0, 1]$, $S = S_0$, $(P, Q, \hat{P}, \hat{Q}) = (P_0, Q_0, \hat{P}_0, \hat{Q}_0)$.

Step 2:   Let loop = loop + 1. If loop = 1, then go to step 4.

Step 3:   Advance the homotopy parameter λ and predict the corresponding $P(\lambda)$, $Q(\lambda)$, $\hat{P}(\lambda)$, and $\hat{Q}(\lambda)$ as follows:

3a: Let $\lambda_0 = \lambda$.

3b: Let $\lambda = \lambda_0 + \Delta\lambda$.

3c: Compute $\dot{P}(\lambda_0)$, $\dot{Q}(\lambda_0)$, $\dot{\hat{P}}(\lambda_0)$, and $\dot{\hat{Q}}(\lambda_0)$ using Eqs. (10-13).

3d: If loop = 2, predict $P(\lambda)$, $Q(\lambda)$, $\hat{P}(\lambda)$, and $\hat{Q}(\lambda)$ using linear prediction, or else predict $P(\lambda)$, $Q(\lambda)$, $\hat{P}(\lambda)$, and $\hat{Q}(\lambda)$ using cubic spline prediction.

3e: Compute the errors $(E_P, E_Q, E_{\hat{P}}, E_{\hat{Q}})$ in the maximum entropy equations (1-4). If the max $(\|E_P\|, \|E_Q\|, \|E_{\hat{P}}\|, \|E_{\hat{Q}}\|)$ satisfies some preassigned tolerance, then continue. Otherwise reduce $\Delta\lambda$ and go to step 3b.

Step 4:   Correct the current approximations $P(\lambda)$, $Q(\lambda)$, $\hat{P}(\lambda)$, and $\hat{Q}(\lambda)$ as follows.

4a: Compute the errors $(E_P, E_Q, E_{\hat{P}}, E_{\hat{Q}})$ in the maximum entropy equations (1-4).

4b: Solve Eqs. (22-25) for ΔP, ΔQ, $\Delta\hat{P}$, and $\Delta\hat{Q}$.

4c: Let

$$P(\lambda) \leftarrow P(\lambda) + \Delta P, \qquad Q(\lambda) \leftarrow Q(\lambda) + \Delta Q$$

$$\hat{P}(\lambda) \leftarrow \hat{P}(\lambda) + \Delta\hat{P}, \qquad \hat{Q}(\lambda) \leftarrow \hat{Q}(\lambda) + \Delta\hat{Q}$$

4d: Recompute the errors $(E_P, E_Q, E_{\hat{P}}, E_{\hat{Q}})$ in the maximum entropy equations (1-4). If the max $(\|E_P\|, \|E_Q\|, \|E_{\hat{P}}\|, \|E_{\hat{Q}}\|)$ satisfies some preassigned tolerance, then continue. Otherwise go to step 4b.

Step 5:   If $\lambda = 1$, then stop. Otherwise go to step 2.

Remark 3.   Since the corrections of step 4 correspond to Newton corrections, quadratic convergence can be insured by choosing the prediction tolerance, used in step 3e, sufficiently small. This insures that along the homotopy curve the approximation to $(P(\lambda), Q(\lambda), \hat{P}(\lambda), \hat{Q}(\lambda))$ is close to the optimal value $(P^*(\lambda), Q^*(\lambda), \hat{P}^*(\lambda), \hat{Q}^*(\lambda))$. Hence, the quadratic convergence properties of Newton's method[18] can be realized. This quadratic convergence has been observed in numerous examples.

Remark 4.   The previous homotopy algorithm for maximum entropy design advanced the P and Q equations separately from the $\hat{P}$ and $\hat{Q}$ equations. That is, $P(\lambda)$ and $Q(\lambda)$ were corrected with $\hat{P}(\lambda) = \hat{P}_a(\lambda)$ and $\hat{Q}(\lambda) = \hat{Q}_a(\lambda)$ where $\hat{P}_a(\lambda)$ and $\hat{Q}_a(\lambda)$ are approximations. Similarly, $\hat{P}(\lambda)$ and $\hat{Q}(\lambda)$ were corrected with $P(\lambda) = P_a(\lambda)$ and $Q(\lambda) = Q_a(\lambda)$ where $P_a(\lambda)$ and $Q_a(\lambda)$ are approximations. This iterative scheme tends to converge slowly as the uncertainty level is increased and never exhibits quadratic convergence, no matter how small the prediction tolerance.

Notice that the algorithm relies on solving four coupled Lyapunov equations (10-13) or (22-25) at each prediction step or correction iteration. Efficient solution of these equations makes the algorithm feasible for large-scale systems. The current solution procedure is based on diagonalizing the coefficient matrices A_p and A_q of the coupled Lyapunov equations. This is usually possible. However, it is possible that this diag-

Table 1   Run-time statistics of the maximum
entropy homotopy algorithm

Initial β	Final β	Megaflops	Real time, s	Predictions and corrections
0	0.01	1246	609	43
0.01	0.1	1062	519	36
0.1	1	1062	513	36
1	5	1212	617	41

Table 2   Robustness to simultaneous shifts
in the undamped natural frequencies

β	$\Delta\omega_{min}$, rad/s	$\Delta\omega_{max}$, rad/s
0 (= LQG)	− 0.000075	0.0075
0.01	− 0.0037	0.036
0.1	− 0.080	0.69
1	− 1.6	7.1
5	− 15	94

Fig. 1   Bode plot of SISO ACES transfer function.

onalization will be intractable for some points along the homotopy path. In this case, one could randomly perturb the system matrices so that diagonalization is possible. The perturbation is then removed at the end of the homotopy curve. This type of random perturbation is commonly used in "probability one homotopies."[17] An alternative is to embed a numerical conditioning test in the program to determine whether the coefficient matrices are truly diagonalizable. If they are not, then one can solve the coupled Lyapunov equations using a nondiagonal alternative such as the Schur decomposition.

## V.   Illustration of Maximum Entropy Design Using the ACES Structure

This section illustrates the design of a maximum entropy controller for a 17th-order model of one of the single-input/single-output (SISO) transfer functions of the ACES structure at NASA Marshall Space Flight Center.[21] The actuator and sensor are, respectively, a torque actuator and a collocated rate gyro. The model includes the actuator and sensor dynamics. A first-order all-pass filter was appended to the model to approximate the computational delay associated with digital implementation.

The Bode plots of the open-loop plant are illustrated in Fig. 1. The basic control objective is to provide damping to the lower frequency modes of the structure (i.e., the modes less than 3 Hz) as measured by the rate gyro. The undamped natural frequencies of each of the eight flexible modes are considered uncertain. (Note that there are two modes at 2.4 Hz, one of which is virtually unobservable.) Maximum entropy design is used to add uncertainty to each of these modal frequencies to increase the design robustness. The uncertainty vector $\alpha \in \mathfrak{R}^8$ is given by

$$\alpha = \beta * \alpha_0$$

where each element of $\alpha_0 \in \mathfrak{R}^8$ has unity value, reflecting equal uncertainty in each of the flexible modes and β is a scale factor chosen to represent the level of uncertainty. The precise relationship between β and the allowable frequency perturbations is not currently defined by maximum entropy theory.

For this example, the MATLAB implementation of the maximum entropy homotopy algorithm was run on a 486, 66-MHz personal computer. The only system matrix that was allowed to vary was α; hence, $A_f = A_0$, $B_f = B_0, \ldots, V_{12,f} = V_{12,0}$. Table 1 shows some of the run-time statistics of the program. The highest uncertainty design, corresponding to $\beta = 5$, was obtained in approximately 37 min. Notice that the number of flops and the run time are essentially linear with respect to the log of the scale factor β. This general trend has also been observed in other design examples.
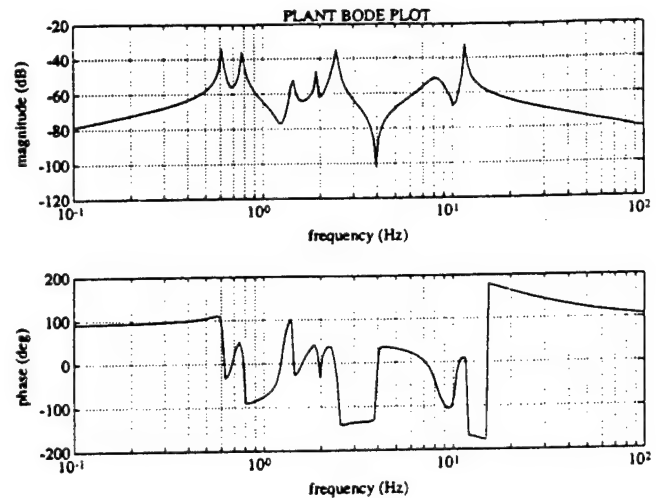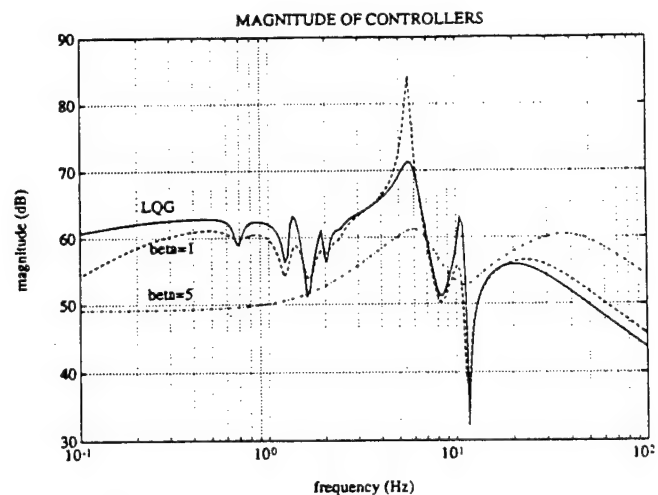
Fig. 2   Magnitude frequency response of LQG and maximum entropy controllers.

As β was increased, the maximum entropy controllers became increasingly more tolerant to changes in the (undamped) natural frequencies. Table 2 describes the robustness properties of the closed-loop systems when the natural frequencies of the open-loop plant were simultaneously shifted by $\Delta\omega$. The parameter $\Delta\omega_{min}$ corresponds to the maximum negative frequency shift, whereas $\Delta\omega_{max}$ corresponds to the maximum positive frequency shift. Notice that the LQG controller is very sensitive to perturbations in the natural frequencies. The maximum entropy controller corresponding to $\beta = 5$ allowed maximum perturbations that were more than four orders of magnitude greater than those allowed by the LQG controller. Robustness analysis that allows independent variations in the modal frequencies can be performed fairly nonconservatively by using theory based on Popov analysis and parameter-dependent Lyapunov functions.[12] An illustration of the application of this theory is given in Ref. 22.

Figures 2 and 3 compare, respectively, the magnitude and phase of the initial LQG controller and the maximum entropy controllers corresponding to $\beta = 1$ and 5. Notice that the $\beta = 5$ controller has a very smooth frequency response and is positive real over a very large frequency band, giving it very significant robustness. The magnitudes of the closed-loop transfer functions corresponding to the LQG compensator and $\beta = 5$ maximum entropy compensator are shown in Fig. 4. As would be expected, the nominal performance (measured by the amount
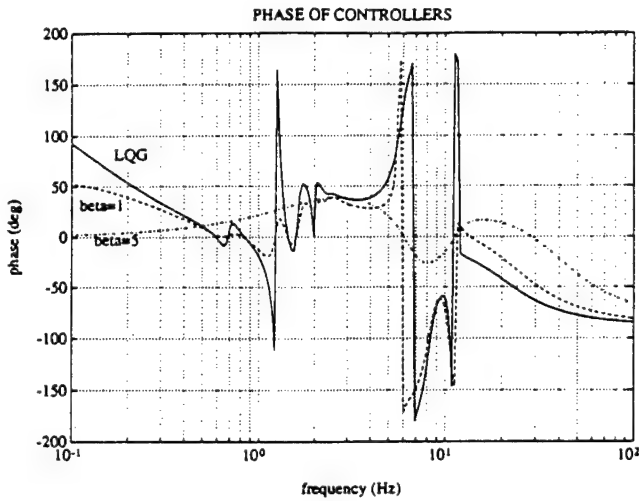
Fig. 3   Phase frequency responses of LQG and maximum entropy controllers.
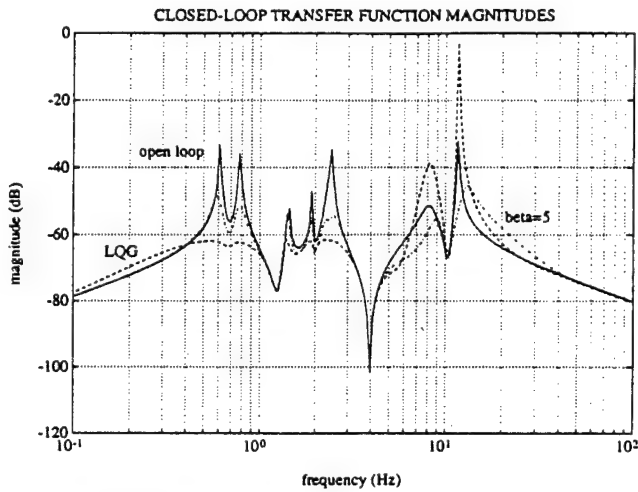

Fig. 4   Magnitude of the closed-loop transfer functions corresponding to the LQG and $\beta = 5$ maximum entropy controller.

of damping in the modes below 2 Hz) of the maximum entropy controller was significantly less than that provided by the LQG controller. However, significant damping was provided by this controller, and as previously discussed, this controller is much more robust than the LQG compensator.

The smoothness of the maximum entropy controller corresponding to $\beta = 5$ indicates that its effective order is much less than 17. Using balanced controller reduction,[23] a fourth-order compensator was obtained whose frequency response is nearly identical to that of the 17th-order compensator. The ability to produce what are essentially reduced-order controllers is an important practical feature of maximum entropy design. Another interesting feature of maximum entropy design is that it will sometimes widen and deepen controller notches to robustly gain stabilize certain modes. This property is illustrated in Refs. 6 and 7. In Ref. 8, maximum entropy design is applied to a multi-input/multi-output control problem, whereas in Ref. 10 maximum entropy design is applied to a neutrally stable system.

VI.   Conclusions

This paper has presented a new homotopy algorithm for maximum entropy control design. The example of the previous section illustrated the use of the algorithm using a medium scale model (17 states) representing a transfer function of the

ACES structure at NASA Marshall Space Flight Center. Very robust designs were obtained in a reasonable amount of time on a 66-MHz, 486 personal computer. For this example, an interesting feature of the most robust maximum entropy controller was that it was essentially a reduced-order controller. This allowed a 17th-order compensator to be easily reduced to a fourth-order compensator by using balanced controller reduction. The frequency responses of the two controllers were essentially identical, indicating that the reduced-order controller maintained the robustness and performance properties of the full-order controller. Algorithms of the type described here should also prove effective in the solution of other sets of coupled Riccati and Lyapunov equations appearing in robust control theory.

Appendix A: Efficient Computation of the Solution to the Maximum Entropy Lyapunov Equation

The Appendix presents a solution procedure for efficiently solving for Q satisfying the $n \times n$ maximum entropy Lyapunov equation

$$0 = A_s Q + Q A_s^T + V + \sum_{i=1}^{n_\alpha} \alpha_i^2 A_i Q A_i^T \qquad (A1)$$

where

$$A_i = e\left(\ell_\alpha(i)\right) e\left(\ell_\alpha(i)+1\right)^T - e\left(\ell_\alpha(i)+1\right) e\left(\ell_\alpha(i)\right)^T \qquad (A2)$$

where $\ell_\alpha \in \mathfrak{R}^{n_\alpha}$ is a vector with distinct elements, each of which lies in the interval $[1 \ n]$, and $e : [1, 2, \ldots, n] \to \mathfrak{R}^n$ is defined by

$$e_i(k) = \begin{cases} 0, & i \neq k \\ 1, & i = k \end{cases}$$

It is assumed that Eq. (A1) has a unique solution. The solution procedure also assumes that A is nondefective and is based on transforming A to a complex, diagonal matrix. Details of the derivation of the solution procedure are given in Ref. 20.

Let Ψ be the eigenvector matrix of A, such that

$$A = \Psi \Lambda \Psi^{-1}$$

where $\Lambda \in \mathbb{C}^{n \times n}$ is diagonal. Then premultiplying and postmultiplying Eq. (A1), respectively, by Ψ^{-1} and Ψ^{-H} yield

$$0 = \Lambda \tilde{Q} + \tilde{Q} \Lambda^* + \bar{V} + M(\tilde{Q})$$

where

$$\tilde{Q} \triangleq \Psi^{-1} Q \Psi^{-H}$$
$$\bar{V} \triangleq \Psi^{-1} V \Psi^{-H}, \qquad M(\tilde{Q}) = \sum_{i=1}^{n_\alpha} M_i(\tilde{Q}) \qquad (A3)$$

and

$$M_i(\tilde{Q}) \triangleq \alpha_i^2 \Psi^{-1} A_i \Psi \tilde{Q} \Psi^H A_i^T \Psi^{-H}$$

The solution procedure relies on the following definitions:

$$\omega_n \triangleq \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \qquad (\omega_n \in \mathfrak{R}^n); \qquad \lambda \triangleq [\Lambda_{11} \ \Lambda_{22} \ \cdots \ \Lambda_{nn}]^T$$

$$S \triangleq -\text{diag}^{-1}(\lambda \omega_n^T + \omega_n \lambda^H) \qquad (A4)$$

$$M_{Q,\alpha} \triangleq \begin{bmatrix} M_{Q,\alpha}^{(1)} & M_{Q,\alpha}^{(2)} & M_{Q,\alpha}^{(3)} \end{bmatrix} \qquad (A5)$$

where

$$M_{Q,\alpha}^{(1)} = \left(\omega_n \otimes \Psi^{-1}(:,\ell_\alpha)\right) * \left(\Psi^{-*}(:,\ell_\alpha) \otimes \omega_n\right)$$

$$M_{Q,\alpha}^{(2)} = \left(\omega_n \otimes \Psi^{-1}(:,\ell_\alpha + \omega_n)\right) * \left(\Psi^{-*}(:,\ell_\alpha + \omega_{n_\alpha}) \otimes \omega_n\right)$$

$$M_{Q,\alpha}^{(3)} = \left(\omega_n \otimes \Psi^{-1}(:,\ell_\alpha)\right) * \left(\Psi^{-*}(:,\ell_\alpha + \omega_{n_\alpha}) \otimes \omega_n\right)$$

$$+ \left(\omega_n \otimes \Psi^{-1}(:,\ell_\alpha + \omega_{n_\alpha})\right) * \left(\Psi^{-*}(:,\ell_\alpha) \otimes \omega_n\right)$$

where

$$N_{Q,\alpha} \triangleq \begin{bmatrix} N_{Q,\alpha}^{(1)} \\ N_{Q,\alpha}^{(2)} \\ N_{Q,\alpha}^{(3)} \end{bmatrix} \qquad (A6)$$

where

$$N_{Q,\alpha}^{(1)} = \left((\alpha * \alpha)\omega_{n2}^T\right)$$

$$* \left[\left(\Psi^*(\ell_\alpha + \omega_{n_\alpha}, :) \otimes \omega_n^T\right) + \left(\omega_n^T \otimes \Psi(\ell_\alpha + \omega_{n_\alpha}, :)\right)\right]$$

$$N_{Q,\alpha}^{(2)} = \left((\alpha * \alpha)\omega_{n2}^T\right)$$

$$* \left[\left(\Psi^*(\ell_\alpha, :) \otimes \omega_n^T\right) + \left(\omega_n^T \otimes \Psi(\ell_\alpha, :)\right)\right]$$

$$N_{Q,\alpha}^{(3)} = -\left((\alpha * \alpha)\omega_{n2}^T\right)$$

$$* \left[\left(\Psi^*(\ell_\alpha + \omega_{n_\alpha}, :) \otimes \omega_n^T\right) + \left(\omega_n^T \otimes \Psi(\ell_\alpha, :)\right)\right]$$

$$P_{Q,\alpha} \triangleq (I - N_{Q,\alpha} S \, M_{Q,\alpha})^{-1} N_{Q,\alpha} \qquad (A7)$$

$$T_Q \triangleq S \, M_{Q,\alpha} P_{Q,\alpha} + I_n \qquad (A8)$$

Summary of Solution Procedure

Step 1: Compute S, $M_{Q,\alpha}$, and $N_{Q,\alpha}$ satisfying, respectively, Eqs. (A4–A6).
Step 2: Compute $P_{Q,\alpha}$ satisfying Eq. (A7).
Step 3: Compute T_Q satisfying Eq. (A8).
Step 4: Compute $\tilde{Q}$ satisfying

$$\text{vec}(\tilde{Q}) = T_Q S \, \text{vec}(\bar{V})$$

Step 5: Compute Q satisfying Eq. (A3) or equivalently

$$Q = \Psi \tilde{Q} \Psi^H$$

Remark A.1. An intermediate step in the derivation of the solution procedure is that

$$\text{vec}\left(M(\tilde{Q})\right) = M_{Q,\alpha} z(\tilde{Q})$$

where

$$z(\tilde{Q}) \triangleq \begin{bmatrix} z_{11}(\tilde{Q}) \\ z_{22}(\tilde{Q}) \\ z_{12}(\tilde{Q}) \end{bmatrix}$$

and

$$z_{11,i}(\tilde{Q}) \triangleq \alpha_i^2 \Psi(k+1), :)\tilde{Q}\Psi^H(:, k+1)$$

$$z_{22,i}(\tilde{Q}) \triangleq \alpha_i^2 \Psi(k, :)\tilde{Q}\Psi^H(:, k)$$

$$z_{12,i}(\tilde{Q}) \triangleq -\alpha_i^2 \Psi(k+1, :)\tilde{Q}\Psi^H(k, :)$$

Appendix B: Efficient Computation of the Solution to Four Coupled Lyapunov Equations for Differentiation and Correction

This Appendix develops a solution procedure for efficiently solving for P, Q, $\hat{P}$, and $\hat{Q}$ satisfying the four $n \times n$ coupled Lyapunov equations

$$0 = A_P^T + PA_P + R + \sum_{i=1}^{n_\alpha} \alpha_i^2 A_i^T PA_i + \sum_{i=1}^{n_\alpha} \alpha_i^2 A_i^T \hat{P}A_i \quad (B1)$$

$$0 = A_Q Q + QA_Q^T + V + \sum_{i=1}^{n_\alpha} \alpha_i^2 A_i Q A_i^T + \sum_{i=1}^{n_\alpha} \alpha_i^2 A_i \hat{Q} A_i^T \quad (B2)$$

$$0 = A_Q^T \hat{P} + \hat{P}A_Q + \hat{R} + G_C Q\hat{F} + \hat{F}QG_C + H_P^H P + PH_P \quad (B3)$$

$$0 = A_P \hat{Q} + \hat{Q}A_Q^T + \hat{V} + G_B P\hat{E} + \hat{E}PG_B + H_Q Q + QH_Q^H \quad (B4)$$

where A_i is defined by Eq. (A2). It is assumed that Eqs. (B1–B4) have a unique solution $(P, Q, \hat{P}, \hat{Q})$. It is also assumed that A_P and A_Q are nondefective. The solution procedure is based on transforming A_P and A_Q to complex diagonal matrices. The results of Appendix A are used extensively. The actual solution procedure is summarized at the end of this Appendix.

Let Ψ_P and Ψ_Q be the eigenvector matrix of A_P and A_Q, such that

$$A_P = \Psi_P \Lambda_P \Psi_P^{-1}, \qquad A_Q = \Psi_Q \Lambda_Q \Psi_Q^{-1} \quad (B5)$$

where $\Lambda_P \in \mathbb{C}^{n \times n}$ and $\Lambda_Q \in \mathbb{C}^{n \times n}$ are diagonal. Substituting Eqs. (B5) into Eqs. (B1–B4) yields

$$0 = \Lambda_P^H \bar{P} + \bar{P}\Lambda_P + \bar{R} + M_P(\bar{P}) + \hat{M}_P(\bar{\bar{P}}) \quad (B6)$$

$$0 = \Lambda_Q \bar{Q} + \bar{Q}\Lambda_Q^H + \bar{V} + M_Q(\bar{Q}) + \hat{M}_Q(\bar{\bar{Q}}) \quad (B7)$$

$$0 = \Lambda_Q^H \bar{\bar{P}} + \bar{\bar{P}}\Lambda_Q + \bar{\bar{R}} + \bar{G}_C \bar{Q}\bar{F} + \bar{F}\bar{Q}\bar{G}_C$$

$$+ \bar{H}_P^H \bar{P}\bar{K}_P + \bar{K}_P^H \bar{P}\bar{H}_P \quad (B8)$$

$$0 = \Lambda_P \bar{\bar{Q}} + \bar{\bar{Q}}\Lambda_P^H + \bar{\bar{V}} + \bar{G}_B \bar{P}\bar{E} + \bar{E}\bar{P}\bar{G}_B + \bar{H}_Q \bar{Q}\bar{K}_Q^H$$

$$+ \bar{K}_Q \bar{Q}\bar{H}_Q^H \quad (B9)$$

where

$$\bar{P} = \Psi_P^H P \Psi_P, \qquad \bar{Q} = \Psi_Q^{-1} Q \Psi_Q^{-H} \quad (B10)$$

$$\bar{\bar{P}} = \Psi_Q^H \hat{P} \Psi_Q, \qquad \bar{\bar{Q}} = \Psi_P^{-1} \hat{Q} \Psi_P^{-H} \quad (B11)$$

$$\bar{R} = \Psi_P^H R \Psi_P, \qquad \bar{V} = \Psi_Q^{-1} V \Psi_Q^{-H}$$

$$\bar{\bar{R}} = \Psi_Q^H \hat{R} \Psi_Q, \qquad \bar{\bar{V}} = \Psi_P^{-1} \hat{V} \Psi_P^{-H}$$

$$M_P(\bar{\bar{P}}) = \sum_{i=1}^{n_\alpha} \alpha_i^2 \Psi_P^H A_i^T \Psi_P^{-H} \bar{P} \Psi_P^{-1} A_i \Psi_P$$

$$\hat{M}_P(\bar{P}) = \sum_{i=1}^{n_\alpha} \alpha_i^2 \Psi_P^H A_i^T \Psi_Q^{-H} \bar{\bar{P}} \Psi_Q^{-1} A_i \Psi_P$$

$$M_Q(\bar{Q}) = \sum_{i=1}^{n_\alpha} \alpha_i^2 \Psi_Q^{-1} A_i \Psi_Q \bar{Q} \Psi_Q^H A_i^T \Psi_Q^{-H}$$

$$\hat{M}_Q(\bar{\bar{Q}}) = \sum_{i=1}^{n_\alpha} \alpha_i^2 \Psi_Q^{-1} A_i \Psi_P \bar{\bar{Q}} \Psi_P^H A_i^T \Psi_Q^{-H}$$

$$\bar{G}_C = \Psi_Q^H G_C \Psi_Q, \qquad \bar{G}_B = \Psi_P^{-1} G_B \Psi_P^{-H}$$

$$\bar{F} = \Psi_Q^H \hat{F} \Psi_Q, \qquad \bar{E} = \Psi_P^{-1} \hat{E} \Psi_P^{-H}$$

$$\bar{H}_P = \Psi_P^{-1} H_P \Psi_Q, \qquad \bar{H}_Q = \Psi_P^{-1} H_Q \Psi_Q$$

$$\bar{K}_P = \Psi_P^{-1} \Psi_Q, \qquad \bar{K}_Q = \Psi_P^{-1} \Psi_Q$$

For $\lambda \in \mathfrak{R}^n$ and $\Psi \in \mathfrak{R}^{n \times n}$ the functions **seig**, **malpha**, and **nalpha** are defined as follows:

$$S = \mathbf{seig}(\lambda)$$

is equivalent to

$$S = \left[-\operatorname{diag}(\lambda \omega_n^T + \omega_n \lambda^H) \right]^{-1}$$

$$M_\alpha = \mathbf{malpha}(\Psi)$$

is equivalent to

$$M_\alpha = \left[M_\alpha^{(1)} \; M_\alpha^{(2)} \; M_\alpha^{(3)} \right]$$

where

$$M_\alpha^{(1)} = \left(\omega_n \otimes \Psi(:, \ell_\alpha) \right) * \left(\Psi^*(:, \ell_\alpha) \otimes \omega_n \right)$$

$$M_\alpha^{(2)} = \left(\omega_n \otimes \Psi(:, \ell_\alpha + \omega_{n_\alpha}) \right) * \left(\Psi(:, \ell_\alpha + \omega_{n_\alpha}) \otimes \omega_n \right)$$

$$M_\alpha^{(3)} = \left(\omega_n \otimes \Psi(:, \ell_\alpha) \right) * \left(\Psi(:, \ell_\alpha + \omega_{n_\alpha}) \otimes \omega_n \right)$$
$$+ \left(\omega_n \otimes \Psi(:, \ell_\alpha + \omega_{n_\alpha}) \right) * \left(\Psi(:, \ell_\alpha) \otimes \omega_n \right)$$

$$N_\alpha = \mathbf{nalpha}(\Psi)$$

is equivalent to

$$N_\alpha = \begin{bmatrix} N_\alpha^{(1)} \\ N_\alpha^{(2)} \\ N_\alpha^{(3)} \end{bmatrix}$$

where

$$N_\alpha^{(1)} = \left((\alpha * \alpha) \omega_{n2}^T \right) * \left[\left(\Psi^*(\ell_\alpha + \omega_{n_\alpha}, :) \otimes \omega_n^T \right) \right.$$
$$\left. + \left(\omega_n^T \otimes \Psi(\ell_\alpha + \omega_{n_\alpha}, :) \right) \right]$$

$$N_\alpha^{(2)} = \left((\alpha * \alpha) \omega_{n2}^T \right) * \left[\left(\Psi^*(\ell_\alpha, :) \otimes \omega_n^T \right) + \left(\omega_n^T \otimes \Psi(\ell_\alpha, :) \right) \right]$$

$$N_\alpha^{(3)} = - \left((\alpha * \alpha) \omega_{n2}^T \right) * \left[\left(\Psi^*(\ell_\alpha + \omega_{n_\alpha}, :) \otimes \omega_n^T \right) * \left(\omega_n^T \otimes \Psi(\ell_\alpha, :) \right) \right]$$

It follows from the results of Appendix A that Eqs. (B6) and (B7) can be expressed as

$$\operatorname{vec}(\check{P}) = T_P S_P \operatorname{vec}\left(\hat{M}_P(\bar{\bar{P}}) \right) + T_P S_P \operatorname{vec}(\bar{R}) \quad \text{(B12)}$$

$$\operatorname{vec}(\check{Q}) = T_Q S_Q \operatorname{vec}\left(\hat{M}_Q(\bar{\bar{Q}}) \right) + T_Q S_Q \operatorname{vec}(\bar{V}) \quad \text{(B13)}$$

where

$$T_P = (S_P M_{P,\alpha} P_{P,\alpha} + I_n), \qquad T_Q = (S_Q M_{Q,\alpha} Q_{Q,\alpha} + I_n)$$

$$P_{P,\alpha} = (I_n - N_{P,\alpha} S_P M_{P,\alpha})^{-1} N_{P,\alpha}$$

$$Q_{Q,\alpha} = (I_n - N_{Q,\alpha} S_Q M_{Q,\alpha})^{-1} N_{Q,\alpha}$$

$$S_P = \mathbf{seig}(\lambda_P^*), \qquad S_Q = \mathbf{seig}(\lambda_Q)$$

$$\lambda_P = [\Lambda_{P,11} \; \Lambda_{P,22} \; \cdots \; \Lambda_{P,nn}]^T$$

$$\lambda_Q = [\Lambda_{Q,11} \; \Lambda_{Q,22} \; \cdots \; \Lambda_{Q,nn}]^T$$

$$M_{P,\alpha} = \mathbf{malpha}(\Psi_P^*), \qquad M_{Q,\alpha} = \mathbf{malpha}(\Psi_Q^{-1})$$

$$N_{P,\alpha} = \mathbf{nalpha}(\Psi_P^{-*}), \qquad N_{Q,\alpha} = \mathbf{nalpha}(\Psi_Q)$$

Using standard Kronecker algebra, we can express Eqs. (B8) and (B9) as

$$\operatorname{vec}(\bar{\bar{P}}) = S_Q^* U_{P,1} \operatorname{vec}(\check{P}) + S_Q^* U_{Q,1} \operatorname{vec}(\check{Q}) + S_Q^* \operatorname{vec}(\bar{\bar{R}}) \quad \text{(B14)}$$

$$\operatorname{vec}(\bar{\bar{Q}}) = S_P^* U_{P,2} \operatorname{vec}(\check{P}) + S_P^* U_{Q,2} \operatorname{vec}(\check{Q}) + S_P^* \operatorname{vec}(\bar{\bar{V}}) \quad \text{(B15)}$$

where

$$U_{P,1} = \left(K_P^T \otimes \bar{H}_P^H \right) + \left(\bar{H}_P^T \otimes K_P^H \right)$$
$$U_{Q,1} = \left(\bar{F}^T \otimes \bar{G}_C \right) + \left(\bar{G}_C^T \otimes \bar{F} \right) \quad \text{(B16)}$$

$$U_{P,2} = \left(\bar{E}^T \otimes \bar{G}_B \right) + \left(\bar{G}_B^T \otimes \bar{E} \right)$$
$$U_{Q,2} = (\bar{K}_Q^* \otimes \bar{H}_Q) + (\bar{H}_Q^* \otimes \bar{K}_Q) \quad \text{(B17)}$$

Now, from the results of Appendix A, we can write

$$\operatorname{vec}\left(\hat{M}_P(\bar{\bar{P}}) \right) = M_{P,\alpha} z(\bar{\bar{P}}), \qquad \operatorname{vec}\left(\hat{M}_Q(\bar{\bar{Q}}) \right) = M_{Q,\alpha} z(\bar{\bar{Q}}) \quad \text{(B18)}$$

$$z(\bar{\bar{P}}) = \hat{N}_{P,\alpha} \operatorname{vec}(\bar{\bar{P}}), \qquad z(\bar{\bar{Q}}) = \hat{N}_{Q,\alpha} \operatorname{vec}(\bar{\bar{Q}}) \quad \text{(B19)}$$

where

$$\hat{N}_{P,\alpha} = \mathbf{nalpha}(\Psi_Q^{-*}), \qquad \hat{N}_{Q,\alpha} = \mathbf{nalpha}(\Psi_P)$$

Substituting, Eqs. (B12) and (B13) into Eqs. (B14) and (B15) gives

$$\operatorname{vec}(\bar{\bar{P}}) = S_Q^* U_{P,1} T_P S_P \operatorname{vec}\left(\hat{M}_P(\bar{\bar{P}}) \right)$$
$$+ S_Q^* U_{Q,1} T_Q S_Q \operatorname{vec}\left(\hat{M}_Q(\bar{\bar{Q}}) \right) + \hat{p}_0 \quad \text{(B20)}$$

$$\operatorname{vec}(\bar{\bar{Q}}) = S_P^* U_{P,2} T_P S_P \operatorname{vec}\left(\hat{M}_P(\bar{\bar{P}}) \right)$$
$$+ S_P^* U_{Q,2} T_Q S_Q \left(\hat{M}_Q(\bar{\bar{Q}}) \right) + \hat{q}_0 \quad \text{(B21)}$$

where

$$\hat{p}_0 = S_Q^* U_{P,1} T_P S_P \operatorname{vec}(\bar{R}) + S_Q^* U_{Q,1} T_Q S_Q \operatorname{vec}(\bar{V}) + S_Q^* \operatorname{vec}(\bar{\bar{R}}) \quad \text{(B22)}$$

$$\hat{q}_0 = S_P^* U_{P,2} T_P S_P \operatorname{vec}(\bar{R}) + S_P^* U_{Q,2} T_Q S_Q \operatorname{vec}(\bar{V}) + S_P^* \operatorname{vec}(\bar{\bar{V}}) \quad \text{(B23)}$$

Substituting Eqs. (B18) into Eqs. (B20) and (B21) gives

$$\operatorname{vec}(\bar{\bar{P}}) = S_Q^* U_{P,1} T_P S_P M_{P,\alpha} z(\bar{\bar{P}})$$
$$+ S_Q^* U_{Q,1} T_Q S_Q M_{Q,\alpha} z(\bar{\bar{Q}}) + \hat{p}_0 \quad \text{(B24)}$$

$$\operatorname{vec}(\bar{\bar{Q}}) = S_P^* U_{P,2} T_P S_P M_{P,\alpha} z(\bar{\bar{P}})$$
$$+ S_P^* U_{Q,2} T_Q S_Q M_{Q,\alpha} z(\bar{\bar{Q}}) + \hat{q}_0 \quad \text{(B25)}$$

Substituting Eqs. (B24) and (B25) into Eq. (B19) gives

$$z(\bar{\bar{P}}) = \hat{N}_{P,\alpha} S_Q^* U_{P,1} T_P S_P M_{P,\alpha} z(\bar{\bar{P}})$$
$$+ \hat{N}_{P,\alpha} S_Q^* U_{Q,1} T_Q S_Q M_{Q,\alpha} z(\bar{\bar{Q}}) + \hat{N}_{P,\alpha} \hat{p}_0$$

$$z(\bar{\bar{Q}}) = \hat{N}_{Q,\alpha} S_P^* U_{P,2} T_P S_P M_{P,\alpha} z(\bar{\bar{P}})$$
$$+ \hat{N}_{Q,\alpha} S_P^* U_{Q,2} T_Q S_Q M_{Q,\alpha} z(\bar{\bar{Q}}) + \hat{N}_{Q,\alpha} \hat{q}_0$$

or, equivalently,

$$\begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \begin{bmatrix} z(\bar{P}) \\ z(\bar{Q}) \end{bmatrix} = \begin{bmatrix} \hat{N}_{P,\alpha}\hat{p}_0 \\ \hat{N}_{Q,\alpha}\hat{q}_0 \end{bmatrix} \qquad \text{(B26)}$$

where

$$D_{11} = I_{3n_\alpha} - \hat{N}_{P,\alpha}S_Q^* U_{P,1}T_P S_P M_{P,\alpha}$$

$$D_{12} = -\hat{N}_{P,\alpha}S_Q^* U_{Q,1}T_Q S_Q M_{Q,\alpha} \qquad \text{(B27)}$$

$$D_{21} = -\hat{N}_{Q,\alpha}S_P^* U_{P,2}T_P S_P M_{P,\alpha}$$

$$D_{22} = I_{3n_\alpha} - \hat{N}_{Q,\alpha}S_P^* U_{Q,2}T_Q S_Q M_{Q,\alpha} \qquad \text{(B28)}$$

Finally, substituting Eqs. (B18) into Eqs. (B12) and (B13) gives

$$\text{vec}(\bar{P}) = T_P S_P M_{P,\alpha} z(\bar{P}) + T_P S_P \text{vec}(\hat{R}) \qquad \text{(B29)}$$

$$\text{vec}(\bar{Q}) = T_Q S_Q M_{Q,\alpha} z(\bar{Q}) + T_Q S_Q \text{vec}(\bar{V}) \qquad \text{(B30)}$$

Notice from Eqs. (B16) and (B17) that $U_{P,1}$, $U_{Q,1}$, $U_{P,2}$, and $U_{Q,2}$ are each an $n^2 \times n^2$ matrix. The storage required to compute these matrices is hence very large for large n. To avoid this memory requirement it is possible to compute $\hat{p}_0$ and $\hat{q}_0$ satisfying Eqs. (B22) and (B23) and D_{11}, D_{12}, D_{21}, and D_{22} satisfying Eqs. (B27) using the identity

$$\text{vec}(ADB) = (B^T \otimes A)\text{vec}(D)$$

By substituting Eqs. (B16) and (B17) into Eqs. (B19), (B27), and (B28), and using Eq. (B29), it follows that $\hat{p}_0$, $\hat{q}_0$, D_{11}, D_{12}, D_{21}, and D_{22} can be computed using the following algorithms. In these algorithms $\text{vec}_n^{-1} : \mathfrak{R}^n \to \mathfrak{R}^{n \times n}$ is understood to be the operator satisfying

$$M = \text{vec}_n^{-1}\left(\text{vec}(M)\right)$$

Algorithm for computation of $\hat{p}_0$ and $\hat{q}_0$:

$$W_P = \text{vec}_n^{-1}\left((M_{P,\alpha}P_{P,\alpha} + I_n)S_P \text{vec}(R)\right)$$

$$W_Q = \text{vec}_n^{-1}\left((M_{Q,\alpha}P_{Q,\alpha} + I_n)S_Q \text{vec}(V)\right)$$

$$\hat{p}_0 = \text{vec}\left((G_c W_Q \hat{F} + H_P^H W_P K_P) + (G_c W_Q \hat{F} + H_P^H W_P K_P)^H + \hat{R}\right)$$

$$\hat{q}_0 = \text{vec}\left((G_B W_P \hat{E} + H_Q W_Q K_Q^H) + (G_B W_P \hat{E} + H_Q W_Q K_Q^H)^H + \hat{V}\right)$$

Algorithm for computation of D_{11}, D_{12}, D_{21}, and D_{22}:

$$V_P = T_P S_P M_{P,\alpha}, \qquad V_Q = T_Q S_Q M_{Q,\alpha}$$

for $i = 1 : 3n_\alpha$

$$D_{11}(:,i) = I_{3n_\alpha} - \hat{N}_{P,\alpha}S_Q^* \text{vec}\left(H_P^H V_P(:,i)K_P + K_P^H V_P(:,i)^H H_P\right)$$

$$D_{12}(:,i) = -\hat{N}_{P,\alpha}S_Q^* \text{vec}\left(G_c V_Q(:,i)\hat{F} + \hat{F}^H V_Q(:,i)^H G_c^H\right)$$

$$D_{21}(:,i) = -\hat{N}_{Q,\alpha}S_P^* \text{vec}\left(H_Q V_P(:,i)\hat{K}_Q^H + K_Q V_P(:,i)^H H_Q^H\right)$$

$$D_{22}(:,i) = I_{3n_\alpha} - \hat{N}_{Q,\alpha}S_P^*\left(H_Q V_Q(:,i)K_Q^H + K_Q V_Q(:,i)^H H_Q^H\right)$$

Summary of Solution Procedure

Step 1: Construct D_{11}, D_{12}, D_{21}, and D_{22} and solve Eq. (B26) for $z(\bar{P})$ and $z(\bar{Q})$.

Step 2: Solve Eqs. (B29) and (B30) for $\bar{P}$ and $\bar{Q}$.

Step 3: Solve Eqs. (B24) and (B25) for $\hat{P}$ and $\hat{Q}$.

Step 4: Compute P, Q, $\hat{P}$, and $\hat{Q}$, satisfying Eqs. (B10) and (B11), or equivalently

$$P = \Psi_P^{-H}\hat{P}\Psi_P^{-1}, \qquad Q = \Psi_Q \hat{Q}\Psi_Q^H$$

$$\hat{Q} = \Psi_Q^{-H}\bar{P}\Psi_Q^{-1}, \qquad \hat{Q} = \Psi_P \bar{Q}\Psi_P^H$$

References

[1]Kwakernaak, H., and Sivan, R., *Linear Optimal Control Systems*, Wiley, New York, 1972.

[2]Doyle, J. C., "Guaranteed Margins for LQG Regulators," *IEEE Transactions on Automatic Control*, Vol. 23, Aug. 1978, pp. 756, 757.

[3]Hyland, D. C., "Maximum Entropy Stochastic Approach to Controller Design for Uncertain Structural Systems," *Proceedings of the American Control Conference* (Arlington, VA), IEEE, Piscataway, NJ, 1982, pp. 680–688.

[4]Bernstein, D. S., and Hyland, D. C., "The Optimal Projection/Maximum Entropy Approach to Designing Low-Order, Robust Controllers for Flexible Structures," *Proceedings of the IEEE Conference on Decision and Control* (Fort Lauderdale, FL), IEEE, Piscataway, NJ, 1985, pp. 745–752.

[5]Bernstein, D. S., and Hyland, D. C., "The Optimal Projection Approach to Robust, Fixed-Structure Control Design," *Mechanics and Control of Large Flexible Structures*, edited by J. L. Junkins, AIAA, Washington, DC, 1990, pp. 237–293.

[6]Collins, E. G., Jr., Phillips, D. J., and Hyland, D. C., "Robust Decentralized Control Laws for the ACES Structure," *Control Systems Magazine*, Vol. 11, April 1991, pp. 62–70.

[7]Collins, E. G., Jr., King, J. A., Phillips, D. J., and Hyland, D. C., "High Performance, Accelerometer-Based Control of the Mini-MAST Structure," *Journal of Guidance, Control, and Dynamics*, Vol. 15, No. 4, 1992, pp. 885–892.

[8]Davis, L. D., Hyland, D. C., and Bernstein, D. S., "Application of the Maximum Entropy Design Approach to the Space Control Laboratory Experiment (SCOLE)," Final Rept. for NASA Contract NAS1-17741, NASA Langley Research Center, Langley, VA, Jan. 1985.

[9]Cheung, M.-F., and Yurkovich, S., "On the Robustness of MEOP Design Versus Asymptotic LQG Synthesis," *IEEE Transactions on Automatic Control*, Vol. 33, Nov. 1988, pp. 1061–1065.

[10]Collins, E. G., Jr., King, J. A., and Bernstein, D. S., "Application of Maximum Entropy/Optimal Projection Design Synthesis to a Benchmark Problem," *Journal of Guidance, Control, and Dynamics*, Vol. 15, No. 5, 1992, pp. 1094–1102.

[11]Bernstein, D. S., Haddad, W. M., Hyland, D. C., and Tyan, F., "A Maximum Entropy-Type Lyapunov Function for Robust Stability and Performance Analysis," *Systems and Control Letters*, Vol. 12, 1993, pp. 73–87.

[12]Haddad, W. M., and Bernstein, D. S., "Parameter-Dependent Lyapunov Functions, Constant Real Parameter Uncertainty, and the Popov Criterion in Robust Analysis and Synthesis: Part 1, Part 2," *Proceedings of the IEEE Conference on Decision and Control* (Brighton, England, UK), IEEE, Piscataway, NJ, 1991, pp. 2274–2279, 2617–2623.

[13]Collins, E. G., and Richter, S., "A Homotopy Algorithm for Synthesizing Robust Controllers for Flexible Structures Via the Maximum Entropy Design Equations," *Third Air Force/NASA Symposium on Recent Advances in Multidisciplinary Analysis and Optimization* (San Diego, CA), May 1990, pp. 1449–1454.

[14]Brewer, J. W., "Kronecker Products and Matrix Calculus in System Theory," *IEEE Transactions on Circuit and Systems*, Vol. CAS-25, No. 9, 1978, pp. 772–781.

[15]Garcia, C. B., and Zangwill, W. I., *Pathways to Solutions, Fixed Points and Equilibria*, Prentice-Hall, Englewood Cliffs, NJ, 1981.

[16]Richter, S. L., and DeCarlo, R. A., "Continuation Methods: Theory and Applications," *IEEE Transactions on Circuits and Systems*, Vol. CAS-30, No. 6, 1983, pp. 347–352.

[17]Watson, L. T., "ALGORITHM 652 HOMPACK: A Suite of Codes for Globally Convergent Homotopy Algorithms," *ACM Transactions on Mathematical Software*, Vol. 13, Sept. 1987, pp. 281–310.

[18]Fletcher, R., *Practical Methods of Optimization*, Wiley, New York, 1987.

[19]Richter, S., Davis, L. D., and Collins, E. G., Jr., "Efficient

Computation of the Solutions to Modified Lyapunov Equations,'' *SIAM Journal of Matrix Analysis and Applications*, Vol. 14, No. 2, 1993, pp. 420–434.

[20]Collins, E. G., Jr., Davis, L. D., and Richter, S., ''Homotopy Algorithms for H_2 Optimal Reduced-Order Dynamic Compensation and Maximum Entropy Control,'' Final Rept. for Contract 54-7609, Sandia National Lab., Albuquerque, NM, Aug. 1992.

[21]Irwin, R. D., Jones, V. L., Rice, S. A., Seltzer, S. M., and Tollison, D. J., ''Active Control Technique Evaluation for Spacecraft (ACES),'' Final Rept. to Flight Dynamics Lab. of Wright Aeronautical Labs, AFWAL-TR-88-3038, Wright-Patterson AFB, OH, June 1988.

[22]Collins, E. G., Jr., Haddad, W. M., and Davis, L. D., ''Riccati Equation Approaches for Small Gain, Positivity, and Popov Robustness Analysis,'' *Journal of Guidance, Control, and Dynamics*, Vol. 17, No. 2, 1994, pp. 322–329; also *Proceedings of the 1993 American Control Conference*, IEEE, Piscataway, NJ, 1993, pp. 1079–1083.

[23]Yousuff, A., and Skelton, R. E., ''A Note on Balanced Controller Reduction,'' *IEEE Transactions on Automatic Control*, Vol. AC-29, March 1984, pp. 254–257.

Appendix C:

The Multivariable Parabola Criterion for Robust Controller Synthesis:
A Riccati Equation Approach

The Multivariable Parabola Criterion for Robust Controller Synthesis: A Riccati Equation Approach

by

Wassim M. Haddad
Department of Mechanical and
Aerospace Engineering
Florida Institute of Technology
Melbourne, FL 32901
(407) 768-8000 Ext. 7241
(407) 984-8461 (FAX)

Dennis S. Bernstein
Department of Aerospace Engineering
The University of Michigan
Ann Arbor, MI 48109-2140
(313) 764-3719
(313) 763-0578 (FAX)

Abstract

In 1967 Bergen and Sapiro derived an absolute (frequency domain) stability criterion that unifies the classical circle and Popov criteria. A slightly weaker version of this combined criterion has a graphical interpretation in the Popov (rather that Nyquist) plane in terms of a parabola. Our goal in this paper is to generalize the parabola criterion in terms of Riccati equations. Besides poviding a multivariable extension, this formulation clarifies connections to state space bounded real and positive real theory and provides the necessary means for robust controller synthesis.

Key Words: Robust stability and performance, Popov criterion, circle criterion, parameter-dependent Lyapunov functions

Running Title: Multivariable Parabola Criterion

1. Introduction

One of the most basic issues in system theory is the stability of feedback interconnections. Four of the most fundamental results concerning stability of feedback systems are the small gain, positivity, circle, and Popov theorems. In a recent paper [6], each result was specialized to the problem of robust stability involving linear uncertainty, and a Lyapunov function framework was established providing connections between these classical results and robust stability via state space methods. Furthermore, it was pointed out in [6] that both gain and phase properties can be simultaneously accounted for by means of the circle criterion which yields the small gain theorem and positivity theorem as special cases. It is important to note that since positivity theory and bounded real theory can be obtained from the circle criterion and vice versa, all three results can be viewed as equivalent from a mathematical point of view. However, the engineering ramifications of the ability to include phase information can be significant [3]. As shown in [6], the main difference between the small gain, positivity, and circle theorems versus the Popov theorem is that the former results guarantee robustness with respect to arbitrarily, time-varying uncertainty while the latter does not. This is not surprising since the Lyapunov function foundation of the small gain, positivity, and circle theorems is based upon conventional or fixed Lyapunov functions which, of course, guarantee stability with respect to arbitrarily, time-varying perturbations. Since time-varying parameter variations can destabilize a system even when the parameter variations are confined to a region in which constant variations are nondestabilizing, a feedback controller designed for time-varying parameter variations may unnecessarily sacrifice performance when the uncertain real parameters are actually constant.

Whereas the small gain, positivity, and circle results are based upon fixed quadratic Lyapunov functions, the Popov result is based upon a quadratic Lyapunov function that is a *function* of the parametric uncertainty. Thus, in effect, the Popov result guarantees stability by means of a *family* of Lyapunov functions. For robust stability, this situation corresponds to the construction of a parameter-dependent quadratic Lyapunov function [7,8]. A key aspect of this approach (see [7,8]) is the fact that it does *not* apply to arbitrarily time-varying uncertainties, which renders it less conservative than fixed quadratic Lyapunov functions (such as the small gain, positivity, and circle results) in the presence of real, constant parameter uncertainty. A framework for parameter-dependent Lyapunov functions was recently developed in [7,8]. An immediate application of this framework is the reinterpretation and generalization of the classical Popov criterion as a parameter-dependent Lyapunov function for constant linear parametric uncertainty.

1

The main contribution of this paper is the unification of the circle and Popov criteria via a parameter-dependent Lyapunov function framework that yields both results as special cases. The unification of the circle and Popov criteria per se is not new to this paper. Indeed, a parabola test which accomplishes this goal was originally developed in [2] and further studied in [10]. However, these results are confined to SISO systems and rely on graphical techniques. The present paper thus has four specific goals:

1. to provide a general framework for the parabola test in terms of parameter-dependent Lyapunov functions in the spirit of [7,8];

2. to obtain a state space characterization of the parabola test via Riccati equations;

3. to obtain a *multivariable* extension of the parabola test for parametric uncertainty; and

4. to use these results for robust controller synthesis.

To illustrate how the parabola test unifies the circle and Popov criteria, consider the plant G in a feedback configuration with uncertainty block Δ as shown in Figure 1. Introducing the multiplier $I + Ns$ into the loop yields the configuration in Figure 2. Applying positivity to the transfer function $(I + Ns)G$ now yields the familiar Popov test. Next consider the equivalent formulation shown in Figure 3 which involves the introduction of an offset transfer function M_1 inparallel with Δ and in feedback about G. The resulting configuration (Figure 4) now involves a shifted Δ (by M_1) and a bilinear transformation of G. Letting $M_1 = 0$ recovers the Popov formulation while $N = 0$ yields the circle formulation. The simultaneous presence of both N and M_1 leads to the parabola test [2].

Although from a mathematical point of view the use of shifts and bilinear tansformations leads to equivalent results, the use of these transformations can yield less conservative results in practice. In addition, since these transformations do not commute with controller optimization techniques, they must be introduced at an early stage prior to the synthesis procedure.

Notation

$\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r$ real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$

$\mathbb{C}, \mathbb{C}^{r \times s}, \mathbb{C}^r$ complex numbers, $r \times s$ complex matrices, $\mathbb{C}^{r \times 1}$

$\mathbb{E}, \mathrm{tr}, 0_{r \times s}$ expectation, trace, $r \times s$ zero matrix

$\bar{\lambda}$ complex conjugate of $\lambda \in \mathbb{C}$

$I_r, (\ )^{\mathrm{T}}, (\ )^*$ $r \times r$ identity, transpose, complex conjugate transpose

$\mathrm{tr}, \rho(\ ), \sigma_{\max}$ trace, spectral radius, largest singular value

$\mathbb{S}^r, \mathbb{N}^r, \mathbb{P}^r$ $r \times r$ symmetric, nonnegative-definite, positive-definite matrices

$Z_1 \leq Z_2, Z_1 < Z_2$ $Z_2 - Z_1 \in \mathbb{N}^r, Z_2 - Z_1 \in \mathbb{P}^r, Z_1, Z_2 \in \mathbb{S}^r$

$\|Z\|_{\mathrm{F}}$ $[\mathrm{tr}\, ZZ^*]^{1/2}$ (Frobenius matrix norm)

$\|H(s)\|_2$ $[(1/2\pi) \int_{-\infty}^{\infty} \|H(j\omega)\|_{\mathrm{F}}^2 \mathrm{d}\omega]^{1/2}$

2. Robust Stability and Performance Problems

Let $\mathcal{U} \subset \mathbb{R}^{n \times n}$ denote a set of perturbations ΔA of a given nominal dynamics matrix $A \in \mathbb{R}^{n \times n}$. We begin by considering the question of whether or not $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$.

Robust Stability Problem. Determine whether the linear system

$$\dot{x}(t) = (A + \Delta A)x(t), \quad t \in [0, \infty), \tag{2.1}$$

is asymptotically stable for all $\Delta A \in \mathcal{U}$.

To consider the problem of robust performance, we introduce an external disturbance model involving white noise signals as in standard LQG (H_2) theory. The robust performance problem concerns the worst-case H_2 norm, that is, the worst-case (over $\mathcal{U}$) of the expected value of a quadratic form involving outputs $z(t) = Ex(t)$, where $E \in \mathbb{R}^{q \times n}$, when the system is subjected to a standard white noise disturbance $w(t) \in \mathbb{R}^d$ with weighting $D \in \mathbb{R}^{n \times d}$.

Robust Performance Problem. For the disturbed linear system

$$\dot{x}(t) = (A + \Delta A)x(t) + Dw(t), \quad t \in [0, \infty), \tag{2.2}$$

$$z(t) = Ex(t), \tag{2.3}$$

where $w(\cdot)$ is a zero-mean $d-$dimensional white noise signal with intensity I_d, determine a performance bound β satisfying

$$J(\mathcal{U}) \triangleq \sup_{\Delta A \in \mathcal{U}} \limsup_{t \to \infty} \mathbb{E}\{\|z(t)\|_2^2\} \leq \beta. \tag{2.4}$$

As shown in Section 5, (2.2) and (2.3) may denote a control system in closed-loop configuration subjected to external white noise disturbances and for which $z(t)$ denotes the state and control regulation error.

Of course, since D and E may be rank deficient, there may be cases in which a finite performance bound β satisfying (2.4) exists while (2.1) is not asymptotically stable over $\mathcal{U}$. In practice, however, robust performance is mainly of interest when (2.1) is robustly stable. Next, we express the H_2 performance measure (2.4) in terms of the observability Gramian for the pair $(A + \Delta A, E)$. For convenience, define the $n \times n$ nonnegative-definite matrices

$$R \triangleq E^T E, \quad V \triangleq DD^T.$$

4

Lemma 2.1. Suppose $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$. Then

$$J(\mathcal{U}) = \sup_{\Delta A \in \mathcal{U}} \operatorname{tr} P_{\Delta A} V = \sup_{\Delta A \in \mathcal{U}} \|G_{\Delta A}(s)\|_2^2, \tag{2.5}$$

where $P_{\Delta A} \in \mathbb{R}^{n \times n}$ is the unique, nonnegative-definite solution to

$$0 = (A + \Delta A)^T P_{\Delta A} + P_{\Delta A}(A + \Delta A) + R, \tag{2.6}$$

and

$$G_{\Delta A}(s) \triangleq E[sI - (A + \Delta A)]^{-1} D. \tag{2.7}$$

Proof. See [7,8]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

In the present paper our approach is to obtain robust stability as a consequence of sufficient conditions for robust performance. Such conditions are developed in the following sections.

3. Robust Stability and Performance via Parameter-Dependent Lyapunov Functions

The key step in obtaining robust stability and performance is to bound the uncertain terms $\Delta A^T P_{\Delta A} + P_{\Delta A} \Delta A$ in the Lyapunov equation (2.6) by means of a parameter-dependent or adaptive bounding function $\Omega(P, \Delta A)$ which guarantees robust stability by means of a family of Lyapunov functions. As shown in [7,8], this framework corresponds to the construction of a parameter-dependent Lyapunov function that guarantees robust stability. As discussed in [7,8], a key feature of this approach is the fact that it constrains the class of allowable time-varying uncertainties thus reducing conservatism in the presence of constant real parameter uncertainty. The following result is fundamental and forms the basis for all later developments.

Theorem 3.1. Let $\Omega_0 \colon \mathbb{N}^n \to \mathbb{S}^n$ and $P_0 \colon \mathcal{U} \to \mathbb{S}^n$ be such that

$$\Delta A^T P + P \Delta A \le \Omega_0(P) - [A^T P_0(\Delta A) + P_0(\Delta A) A + \Delta A^T P_0(\Delta A) + P_0(\Delta A) \Delta A],$$
$$\Delta A \in \mathcal{U}, P \in \mathbb{N}^n, \tag{3.1}$$

and suppose there exists $P \in \mathbb{N}^n$ satisfying

$$0 = A^T P + P A + \Omega_0(P) + R \tag{3.2}$$

and such that $P + P_0(\Delta A)$ is nonnegative definite for all $\Delta A \in \mathcal{U}$. Then

$$(A + \Delta A, E) \text{ is detectable, } \Delta A \in \mathcal{U}, \tag{3.3}$$

5

if and only if

$$A + \Delta A \text{ is asymptotically stable}, \quad \Delta A \in \mathcal{U}. \tag{3.4}$$

In this case,

$$P_{\Delta A} \leq P + P_0(\Delta A), \quad \Delta A \in \mathcal{U}, \tag{3.5}$$

where $P_{\Delta A}$ is given by (2.9). Therefore,

$$J(\mathcal{U}) \leq \text{tr } PV + \sup_{\Delta A \in \mathcal{U}} \text{tr } P_0(\Delta A)V. \tag{3.6}$$

If, in addition, there exists $\bar{P}_0 \in \mathbf{S}^n$ such that

$$P_0(\Delta A) \leq \bar{P}_0, \quad \Delta A \in \mathcal{U}, \tag{3.7}$$

then

$$J(\mathcal{U}) \leq \beta, \tag{3.8}$$

where

$$\beta \triangleq \text{tr}[(P + \bar{P}_0)V]. \tag{3.9}$$

Proof. We stress that in (3.1) P denotes an arbitrary element of $\mathbf{N}^n$, whereas in (3.2) P denotes a specific solution of the modified Lyapunov equation. This minor abuse of notation considerably simplifies the presentation. To begin, note that for all $\Delta A \in \mathbb{R}^{n \times n}$, (3.2) is equivalent to

$$0 = (A + \Delta A)^\mathrm{T} P + P(A + \Delta A) + \Omega_0(P) - (\Delta A^\mathrm{T} P + P \Delta A) + R. \tag{3.10}$$

Adding and subtracting $A^\mathrm{T} P_0(\Delta A) + P_0(\Delta A)A + \Delta A^\mathrm{T} P_0(\Delta A) + P_0(\Delta A)\Delta A$ to and from (3.10) yields

$$\begin{aligned} 0 =&(A + \Delta A)^\mathrm{T}(P + P_0(\Delta A)) + (P + P_0(\Delta A))(A + \Delta A) \\ &+ \Omega_0(P) - [A^\mathrm{T} P_0(\Delta A) + P_0(\Delta A)A + \Delta A^\mathrm{T} P_0(\Delta A) + P_0(\Delta A)\Delta A] \\ &- (\Delta A^\mathrm{T} P + P \Delta A) + R. \end{aligned} \tag{3.11}$$

Hence, by assumption, (3.11) has a solution $P \in \mathbf{N}^n$ for all $\Delta A \in \mathbb{R}^{n \times n}$. If ΔA is restricted to the set $\mathcal{U}$ then, by (3.1), the expression

$$\Omega_0(P) - [A^\mathrm{T} P_0(\Delta A) + P_0(\Delta A)A + \Delta A^\mathrm{T} P_0(\Delta A) + P_0(\Delta A)\Delta A] - (\Delta A^\mathrm{T} P + P \Delta A)$$

6

is nonnegative definite. Thus, if the detectability condition (3.3) holds for all $\Delta A \in \mathcal{U}$, then it follows from Theorem 3.6 of [12] that $(A + \Delta A, [R + \Omega(P, \Delta A) - (\Delta A^T P + P \Delta A)]^{1/2})$ is detectable for all $\Delta A \in \mathcal{U}$, where

$$\Omega(P, \Delta A) \triangleq \Omega_0(P) - [A^T P_0(\Delta A) + P_0(\Delta A)A + \Delta A^T P_0(\Delta A) + P_0(\Delta A)\Delta A]. \tag{3.12}$$

It now follows from (3.11) and Lemma 12.2 of [12] that $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$. Conversely, if $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$, then (3.3) is immediate. Now, subtracting (2.9) from (3.11) yields

$$
\begin{aligned}
0 = {} & (A + \Delta A)^T(P + P_0(\Delta A) - P_{\Delta A}) + (A + \Delta A)(P + P_0(\Delta A) - P_{\Delta A}) \\
& + \Omega_0(P) - [A^T P_0(\Delta A) + P_0(\Delta A)A + \Delta A^T P_0(\Delta A) + P_0(\Delta A)\Delta A] \\
& - (\Delta A^T P + P \Delta A), \quad \Delta A \in \mathcal{U},
\end{aligned} \tag{3.13}
$$

or, equivalently, since $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$,

$$P + P_0(\Delta A) - P_{\Delta A} = \int_0^\infty e^{(A + \Delta A)^T t} \, \Omega(P, \Delta A) - (\Delta A^T P + P \Delta A)] e^{(A + \Delta A)t} \mathrm{d}t \geq 0, \quad \Delta A \in \mathcal{U}, \tag{3.14}$$

which implies (3.5). The performance bounds (3.6), (3.8) are now an immediate consequence of (2.8), (3.5), and (3.7). $\qquad\square$

Note that, with $\Omega(P, \Delta A)$ defined by (3.12), condition (3.1) can be written as

$$\Delta A^T P + P \Delta A \leq \Omega(P, \Delta A), \quad \Delta A \in \mathcal{U}, \quad P \in \mathbf{N}^n, \tag{3.15}$$

where $\Omega(P, \Delta A)$ is a function of the uncertain parameters ΔA. For convenience we shall say that $\Omega(\cdot, \cdot)$ is a *parameter-dependent bounding function* or, to be consistent with [7,8], a *parameter-dependent Ω-bound*.

Finally, we note that the parameter-dependent Ω-bound framework establishing robust stability given by Theorem 3.1 is equivalent to the existence of a parameter-dependent Lyapunov function of the form

$$V(x) = x^T(P + P_0(\Delta A))x$$

which also establishes robust stability. For further details see [6-8].

4. Construction of Parameter-Dependent Lyapunov Functions and Connections to the Multivariable Parabola Criterion

In this section we assign explicit structure to the set $\mathcal{U}$ and the parameter-dependent bounding function $\Omega(\cdot,\cdot)$. Specifically, the uncertainty set $\mathcal{U}$ is defined by

$$\mathcal{U} \triangleq \{\Delta A \in \mathbb{R}^{n \times n}: \ \Delta A = B_0 F C_0, \text{ where } F \in \mathcal{F}\}, \tag{4.1}$$

where $\mathcal{F}$ is a subset of the set $\hat{\mathcal{F}}$, which is defined by

$$\hat{\mathcal{F}} \triangleq \{F \in \mathbb{R}^{m_0 \times m_0}: (F - M_1)^{\mathrm{T}}[(M_2 - M_1)^{-1} + (M_2 - M_1)^{-\mathrm{T}}](F - M_1) \le (F - M_1) + (F - M_1)^{\mathrm{T}}\}. \tag{4.2}$$

In (4.1) and (4.2), $B_0 \in \mathbb{R}^{n \times m_0}$ and $C_0 \in \mathbb{R}^{m_0 \times n}$ are fixed matrices denoting the structure of the uncertainty, $F \in \mathbb{R}^{m_0 \times m_0}$ is an uncertain matrix, and M_1, M_2 are given $m_0 \times m_0$ matrices such that $(M_2 - M_1)^{-1}$ exists.

Next, we digress slightly to provide simplified characterizations of the set $\mathcal{F}$. Define the subset $\hat{\mathcal{F}}_0$ of $\hat{\mathcal{F}}$ by

$$\hat{\mathcal{F}}_0 = \{F \in \hat{\mathcal{F}}: \ \det[I - (M_2 - M_1)^{-1}(F - M_1)] \ne 0\}. \tag{4.3}$$

Proposition 4.1. The set $\hat{\mathcal{F}}_0$ is equivalently characterized by

$$\hat{\mathcal{F}}_0 = \{F \in \mathbb{R}^{m_0 \times m_0}: \ F = [I + \hat{F}(M_2 - M_1)^{-1}]^{-1}\hat{F} + M_1,$$

$$\text{where } \hat{F} \in \mathbb{R}^{m_0 \times m_0}, \ \hat{F} + \hat{F}^{\mathrm{T}} \ge 0, \text{ and } \det[I + \hat{F}(M_2 - M_1)^{-1}] \ne 0\}.$$

Proof. The proof is an immediate consequence of Proposition 4.1 of [7,8] with F replaced by $F - M_1$. $\qquad\square$

In the special case that $M_2 - M_1$ is positive definite, it follows from Lemma 4.1 of [8] that the condition $\det[I + \hat{F}(M_2 - M_1)^{-1}] \ne 0$ in the definition of $\hat{\mathcal{F}}_0$ is automatically satisfied. In this case, we have the following norm bound inequality on F.

Lemma 4.1. Let $F \in \mathcal{F}$ and assume that $M_2 - M_1 \in \mathbb{P}^{m_0}$. Then

$$\sigma_{\max}(F - M_1) \le \sigma_{\max}(M_2 - M_1). \tag{4.4}$$

Proof. First note that if $F \in \mathcal{F}$, then $M_2 - M_1 \in \mathbb{P}^{m_0}$. In this case it follows from (4.2) that

$$0 \le (F - M_1)^{\mathrm{T}}[2(M_2 - M_1)^{-1}](F - M_1) \le (F - M_1) + (F - M_1)^{\mathrm{T}}.$$

8

Hence, $F \in \mathcal{F}$ implies

$$\lambda_{\max}[(F - M_1)^T Z(F - M_1)] \leq \lambda_{\max}[(F - M_1) + (F - M_1)^T], \tag{4.5}$$

where $Z \triangleq 2(M_2 - M_1)^{-1}$. Now, since $(F - M_1)^T Z(F - M_1)$ and $(F - M_1) + (F - M_1)^T$ are nonnegative definite, (4.5) is equivalent to

$$\sigma_{\max}[(F - M_1)^T Z(F - M_1)] \leq \sigma_{\max}[(F - M_1) + (F - M_1)^T] \leq 2\sigma_{\max}(F - M_1). \tag{4.6}$$

Next, since $Z \in \mathbb{P}^{m_0}$, $\lambda_{\min}(Z)I \leq Z$ or, equivalently, $\sigma_{\min}(Z)I \leq Z$. Hence (4.6) implies

$$\sigma_{\min}(Z)\sigma_{\max}[(F - M_1)^T(F - M_1)] = \sigma_{\max}[(F - M_1)\sigma_{\min}(Z)I(F - M_1)]$$
$$\leq \sigma_{\max}[(F - M_1)^T Z(F - M_1)]$$
$$\leq 2\sigma_{\max}(F - M_1). \tag{4.7}$$

Using $\sigma_{\max}[(F - M_1)^T(F - M_1)] = \sigma_{\max}^2(F - M_1)$, (4.7) yields

$$\sigma_{\min}(Z)\sigma_{\max}^2(F - M_1) \leq 2\sigma_{\max}(F - M_1), \tag{4.8}$$

which proves (4.4). $\qquad\square$

Next, we provide further simplification of the set $\mathcal{F}$ in the case in which $F, M_1,$ and M_2 are symmetric and $M_2 - M_1$ is positive definite.

Lemma 4.2. Let $F, M_1, M_2 \in \mathbb{S}^{m_0}$ and $M_2 - M_1 \in \mathbb{P}^{m_0}$. Then $(F - M_1)(M_2 - M_1)^{-1}(F - M_1) \leq F - M_1$ if and only if $M_1 \leq F \leq M_2$.

Proof. The proof follows as in the proof of Lemma 4.2 of [7,8]. $\qquad\square$

Thus, in the case in which F, M_2, M_1 are symmetric and $M_2 - M_1$ is positive definite, the set $\hat{\mathcal{F}}$ defined by (4.2) becomes

$$\hat{\mathcal{F}}_s \triangleq \{F \in \mathbb{S}^{m_0} : M_1 \leq F \leq M_2\}. \tag{4.9}$$

Note that if F in $\mathcal{F}$ is constrained to have the diagonal structure $\mathrm{diag}[F_1, F_2, \ldots, F_{m_0}]$, then $M_{1i} \leq F_i \leq M_{2i}$, $i = 1, \ldots, m_0$, where $M_1 = \mathrm{diag}[M_{11}, M_{12}, \ldots, M_{1m_0}]$ and $M_2 = \mathrm{diag}[M_{21} M_{22}, \ldots, M_{2m_0}]$. More generally, F may have repeated elements and/or blocks in the diagonal of the form diag $[F_1, F_1, F_1, F_2, \ldots, F_{m_0}]$.

For the structure of $\mathcal{U}$ satisfying (4.1), the parameter-dependent bound $\Omega(\cdot, \cdot)$ satisfying (3.12) can now be given a concrete form. However, since the elements ΔA in $\mathcal{U}$ are parameterized by the

9

elements F in $\mathcal{F}$, for convenience in the following results we shall write $P_0(F)$ in place of $P_0(\Delta A)$. Furthermore, we introduce a key definition that will be used in subsequent developments.

Definition 4.1. Let $M_1, M_2, N \in \mathbb{R}^{m_0 \times m_0}$. Then $\mathcal{F}$ and N are *compatible* if $(F - M_1)^T N$ is symmetric for all $F \in \mathcal{F}$. Furthermore, $\mathcal{F}$ and N are *strongly compatible* if, in addition, $(F - M_1)^T N$ is nonnegative-definite for all $F \in \mathcal{F}$.

Finally, for the remainder of this paper we assume for simplicity that $M_2 - M_1$ is positive definite. In this case it follows from Lemma 4.1 that there exists $\mu \in \mathbb{S}^{m_0}$ such that $(F - M_1)^T N \le \mu$ for all $F \in \mathcal{F}$.

Proposition 4.2. Let $M_1, M_2, N \in \mathbb{R}^{m_0 \times m_0}$ be such that $\mathcal{F}$ and N are compatible and

$$(M_2 - M_1)^{-1} - NC_0 B_0 + [(M_2 - M_1)^{-1} - NC_0 B_0]^T > 0. \tag{4.10}$$

Then the functions

$$\Omega_0(P) = \left[C_0 + NC_0(A + B_0 M_1 C) + B_0^T P\right]^T \left[(M_2 - M_1)^{-1} - NC_0 B_0 + [(M_2 - M_1)^{-1} - NC_0 B_0]^T\right]^{-1}$$
$$\cdot [C_0 + NC_0(A + B_0 M_1 C) + B_0^T P] + PB_0 M_1 C_0 + C_0^T M_1^T B_0^T P, \tag{4.11}$$

$$P_0(F) = C_0^T (F - M_1)^T N C_0, \tag{4.12}$$

satisfy (3.1) with $\mathcal{U}$ given by (4.1).

Proof. Since by (4.3) $(M_2 - M_1)^{-1} - NC_0 B_0 + [(M_2 - M_1)^{-1} - NC_0 B_0]^T > 0$ and by (4.2) $F - M_1 + (F - M_1)^T - (F - M_1)^T [2(M_2 - M_1)^{-1}](F - M_1) \ge 0$ it follows that

$$0 \le \left[[C_0 + NC_0(A + B_0 M_1 C_0) + B_0^T P] - [(M_2 - M_1)^{-1} - NC_0 B_0 + (M_2 - M_1)^{-1} - NC_0 B_0)^T](F - M_1)C_0\right]^T$$
$$\cdot [(M_2 - M_1)^{-1} - NC_0 B_0 + ((M_2 - M_1)^{-1} - NC_0 B_0)^T]^{-1}$$
$$\cdot \left[[C_0 + NC_0(A + B_0 M_1 C_0) + B_0^T P] - [(M_2 - M_1)^{-1} - NC_0 B_0 + ((M_2 - M_1)^{-1} - NC_0 B_0)^T](F - M_1)C_0\right]$$
$$+ C_0^T \left[(F - M_1) + (F - M_1)^T - (F - M_1)^T[2(M_2 - M_1)^{-1}](F - M_1)\right] C_0$$
$$= \Omega_0(P) - PB_0 M_1 C_0 - C_0^T M_1^T B_0^T P - [C_0 + NC_0(A + B_0 M_1 C_0) + B_0^T P]^T (F - M_1)C_0$$
$$- C_0^T (F - M_1)^T[C_0 + NC_0(A + B_0 M_1 C_0) + B_0^T P]$$
$$+ C_0^T (F - M_1)^T[(M_2 - M_1)^{-1} - NC_0 B_0 + ((M_2 - M_1) - NC_0 B_0)^T](F - M_1)C_0$$
$$+ C_0^T[(F - M_1) + (F - M_1)^T - (F - M_1)^T[2(M_2 - M_1)^{-1}](F - M_1)]C_0$$
$$= \Omega_0(P) - A^T C_0^T N^T (F - M_1)C_0 - C_0^T M_1^T B_0^T C_0^T N^T (F - M_1)C_0 - PB_0 F C_0$$

10

$$-C_0^T(F - M_1)^T N C_0 A - C_0^T(F - M_1)^T N C_0 B_0 M_1 C_0 - C_0^T F^T B_0^T P$$
$$- C_0^T(F - M_1)^T N C_0 B_0(F - M_1)C_0 - C_0^T(F - M_1)^T B_0^T C_0^T N^T(F - M_1)C_0$$
$$= \Omega_0(P) - [A^T P_0(F) + P_0(F)A + \Delta A^T P_0(F) + P_0(F)\Delta A] - [\Delta A^T P + P\Delta A],$$

which proves (3.1) with $\mathcal{U}$ given by (4.1). $\qquad \square$

Remark 4.1. Note that by setting $M_1 = 0$, one recovers the parameter-dependent Ω-bound considered in [7,8] which corresponds to a generalized multivariable version of the Popov criterion for linear uncertainty.

Remark 4.2. Note that, unlike the results of [7,8], $P_0(0) = -C_0^T M_1^T N C_0 \neq 0$ and $\Omega_0(P)$ is *not* nonnegative definite. For further discussion on indefinite parameter-dependent Ω-bounds resulting in indefinite Riccati/Lyapunov type equations see [4].

Next, using Theorem 3.1 and Proposition 4.2 we have the following immediate result.

Theorem 4.1. Let $M_1, M_2, N \in \mathbb{R}^{m_0 \times m_0}$ be such that $\mathcal{F}$ and N are strongly compatible and (4.3) is satisfied. Furthermore, suppose there exists a nonnegative-definite matrix P satisfying

$$0 = (A + B_0 M_1 C_0)^T P + P(A + B_0 M_1 C_0)$$
$$+ [C_0 + NC_0(A + B_0 M_1 C_0) + B_0^T P]^T[(M_2 - M_1)^{-1} - NC_0 B_0 + ((M_2 - M_1)^{-1} - NC_0 B_0)^T]^{-1}$$
$$\cdot [C_0 + NC_0(A + B_0 M_1 C_0) + B_0^T P] + R. \tag{4.13}$$

Then

$$(A + \Delta A, E) \text{ is detectable}, \quad \Delta A \in \mathcal{U}, \tag{4.14}$$

if and only if

$$A + \Delta A \text{ is asymptotically stable}, \quad \Delta A \in \mathcal{U}. \tag{4.15}$$

In this case,

$$J(\mathcal{U}) \leq \mathrm{tr}\ PV + \sup_{F \in \mathcal{F}} \mathrm{tr}\ C_0^T(F - M_1)^T N C_0 \leq \mathrm{tr}\ [(P + C_0^T \mu C_0)V]. \tag{4.16}$$

Proof. The result is a direct specialization of Theorem 3.1 using Proposition 4.2. We only note that $P_0(\Delta A)$ now has the form $P_0(F) = C_0^T(F - M_1)^T N C_0$. Since by assuption $(F - M_1)^T N \geq 0$ for all $F \in \mathcal{F}$ it follows that $P + P_0(F)$ is nonnegative definite for all $F \in \mathcal{F}$ as requred by Theorem 3.1. $\qquad \square$

Remark 4.3. The condition that $(F - M_1)^T N = N^T(F - M_1)$, $F \in \mathcal{F}$, represents an intimate relationship between the matrix N and the structure $\mathcal{F}$. For example, if $F = F_1 I_{m_0}$, and $M_1 = M I_{m_0}$, then N can be an arbitrary nonnegative-definite matrix. Alternatively, if $N = N_0 I_{m_0}$, then $F - M_1$ may be nondiagonal. Of course, $F - M_1$ and N may have more intricate structure, for example, they may be block diagonal with commuting blocks situated on the diagonal.

Next, we establish connections between the parameter-dependent bounding function formed by (4.11) and (4.12) and the classical parabola test [2,10]. Furthermore, by exploiting results from positivity theory it is possible to guarantee the existence of a positive-definite solution to (4.13). First, however, we present additional notation and definitions and a key lemma concerning strongly positive real transfer functions. Let

$$G(s) \sim \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]$$

denote a state space realization of a transfer function $G(s)$, that is, $G(s) = C(sI - A)^{-1}B + D$. The notation "$\overset{min}{\sim}$" denotes a minimal realization. Furthermore, an *asymptotically stable transfer function* is a transfer function each of whose poles is in the open left half plane.

A square transfer function $G(s)$ is called *positive real* [1, p. 216] if 1) all poles of $G(s)$ are in the closed left half plane and 2) $G(s) + G^*(s)$ is nonnegative definite for $\text{Re}[s] > 0$. A square transfer function $G(s)$ is called *strictly positive real* [9,11] if 1) $G(s)$ is asymptotically stable and 2) $G(j\omega) + G^*(j\omega)$ is positive definite for all real ω. Finally, a square transfer function $G(s)$ is *strongly positive real* if it is strictly positive real and $D + D^T > 0$, where $D \triangleq G(\infty)$.

Lemma 4.3. Let $G(s) \overset{min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right]$. Then the following statements are equivalent:

i) A is asymptotically stable and $G(s)$ is strongly positive real;

ii) $D + D^T > 0$ and there exist positive-definite matrices P and R such that

$$0 = A^T P + PA + (C - B^T P)^T (D + D^T)^{-1}(C - B^T P) + R. \tag{4.17}$$

Proof. See [5]. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Next, using Lemma 4.3 we obtain a sufficient condition for the existence of a solution to (4.13).

Theorem 4.2. Let $\hat{G}(s) \overset{min}{\sim} \left[\begin{array}{c|c} A + B_0 M_1 C_0 & -B_0 \\ \hline C_0 + N C_0(A + B_0 M_1 C_0) & (M_2 - M_1)^{-1} - N C_0 B_0 \end{array} \right]$. If A is asymptotically stable and $\hat{G}(s)$ is strongly positive real then there exists an $n \times n$ matrix $P > 0$

satisfying (4.13). Conversely, if $2(M_2 - M_1)^{-1} - [NC_0B_0 + NC_0B_0] > 0$ and there exists $P > 0$ satisfying (4.13) for all $R > 0$, then A is asymptotically stable and $\hat{G}(s)$ is strongly positive real.

Proof. The proof is an immediate consequence of Lemma 4.3. $\quad\square$

Next, we show that Theorem 4.1 is a generalization of the classical parabola test [2] for the case in which the loop sector-bounded nonlinearity is used to represent uncertainty. First, however, we provide a generalization of the parabola criterion for multivariable systems with diagonal nonlinearity structure. Specifically, we define the set Φ characterizing a class of sector-bounded memoryless *time $-$ invariant* nonlinearities. Let M_1, M_2 and $M_2 - M_1$ be given positive-definite diagonal matrices and define

$$\Phi \triangleq \left\{ \phi \colon \mathbb{R}^{m_0} \to \mathbb{R}^{m_0} \colon (\phi - M_1 y)^{\mathrm{T}}[(M_2 - M_1)^{-1}(\phi - M_1 y) - y] \leq 0, \ y \in \mathbb{R}^{m_0}, \right.$$

$$\left. \text{and } \phi(y) = [\phi_1(y_1), \phi_2(y_2), \ldots, \phi_{m_0}(y_{m_0})]^{\mathrm{T}} \right\}.$$

Note that for $M_1 = \mathrm{diag}[\underline{m}_1, \underline{m}_2, \ldots, \underline{m}_{m_0}]$ and $M_2 = \mathrm{diag}[\overline{m}_1, \overline{m}_2, \ldots, \overline{m}_{m_0}]$, $\underline{m}_i, \overline{m}_i > 0$, $i = 1, \ldots, m_0$, it follows that each component $\phi_i(y_i)$ of ϕ satisfies

$$\underline{m}_i y_i^2 \leq \phi_i(y_i) y_i \leq \overline{m}_i y_i^2, \ y_i \in \mathbb{R}, \ i = 1, \ldots, m_0.$$

Theorem 4.3. (The Multivariable Parabola Criterion). If there exists a nonnegative-definite diagonal matrix N such that $(M_2 - M_1)^{-1} + (I + Ns)(I + G(s)M_1)^{-1}G(s)$ is strongly positive real, where $G(s) \overset{\min}{\sim} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array} \right]$, then the negative feedback interconnection of $G(s)$ and $\phi(\cdot)$ is asymptotically stable for all $\phi(\cdot) \in \Phi$.

Proof. First note that the negative feedback interconnection of $G(s)$ and $\phi(\cdot)$ has the state-space description

$$\dot{x}(t) = Ax(t) - B\phi(y(t)), \tag{4.18}$$

$$y(t) = Cx(t). \tag{4.19}$$

Now, noting that $[I + G(s)M_1]^{-1}G(s)$ corresponds to a plant $G(s)$ with feedback gain M_1, it follows from feedback interconnection manipulations that a minimal realization for $[I + G(s)M_1]^{-1}G(s)$ is given by

$$[I + G(s)M_1]^{-1}G(s) \overset{\min}{\sim} \left[\begin{array}{c|c} A - BM_1C & B \\ \hline C & 0 \end{array} \right].$$

13

Similarly, noting that $sG(s) \sim \left[\begin{array}{c|c} A & B \\ \hline CA & CB \end{array}\right]$ it follows that $(M_2-M_1)^{-1}+(I+Ns)(I+G(s)M_1)^{-1}G(s)$ has a minimal realization given by

$$\left[\begin{array}{c|c} A-BM_1C & B \\ \hline C+NC(A-BM_1C) & (M_2-M_1)^{-1}+NCB \end{array}\right].$$

Now it follows from Lemma 4.3 that since $(M_2-M_1)^{-1}+(I+Ns)(I+G(s)M_1)^{-1}G(s)$ is strongly positive real there exist positive-definite matrices P and R such that

$$
\begin{aligned}
0 =& (A-BM_1C)^{\mathrm{T}}P + P(A-BM_1C) \\
&+ \left[C+NC(A-BM_1C)-B^{\mathrm{T}}P\right]^{\mathrm{T}}\left[(M_2-M_1)^{-1}+NCB+\left((M_2-M_1)^{-1}+NCB\right)^{\mathrm{T}}\right]^{-1} \\
&\cdot [C+NC(A-BM_1C)-B^{\mathrm{T}}P] + R.
\end{aligned}
\tag{4.20}
$$

Next, for $\phi \in \Phi$ define the Lyapunov function

$$V(x) = x^{\mathrm{T}}Px + 2\sum_{i=1}^{m}\int_{0}^{y_i}[\phi_i(\sigma)-\underline{m}_i\sigma]N_i\mathrm{d}\sigma. \tag{4.21}$$

The coresponding Lyapunov derivative is given by

$$\dot{V}(x) = x^{\mathrm{T}}(A^{\mathrm{T}}P+PA)x - \phi^{\mathrm{T}}B^{\mathrm{T}}Px - x^{\mathrm{T}}PB\phi + 2(\phi-M_1y)^{\mathrm{T}}N\dot{y}. \tag{4.22}$$

Next, using (4.20), noting that $\dot{y} = CAx - CB\phi$, and adding and subtracting $2(\phi-M_1y)^{\mathrm{T}}(M_2-M_1)^{-1}(\phi-M_1y)$, $2(\phi-M_1y)^{\mathrm{T}}y$, $2x^{\mathrm{T}}C^{\mathrm{T}}M_1NCB\phi$, $2x^{\mathrm{T}}A^{\mathrm{T}}C^{\mathrm{T}}NM_1Cx$, $2x^{\mathrm{T}}C^{\mathrm{T}}M_1B^{\mathrm{T}}C^{\mathrm{T}}N\phi$, and $2x^{\mathrm{T}}C^{\mathrm{T}}M_1B^{\mathrm{T}}C^{\mathrm{T}}NM_1Cx$ to and from (4.22) it follows (after some algebraic manipulation) that

$$\dot{V}(x) = -x^{\mathrm{T}}Rx - z^{\mathrm{T}}z + 2(\phi-M_1y)^{\mathrm{T}}[(M_2-M_1)^{-1}(\phi-M_1y)-y],$$

where

$$
\begin{aligned}
z \triangleq & \left[(M_2-M_1)^{-1}+NCB+\left((M_2-M_1)^{-1}+NCB\right)^{\mathrm{T}}\right]^{-1/2}[C+NC(A-BM_1C)-B^{\mathrm{T}}P]x \\
& - \left[(M_2-M_1)^{-1}+NCB+\left((M_2-M_1)^{-1}+NCB\right)^{\mathrm{T}}\right]^{1/2}[\phi-M_1Cx].
\end{aligned}
$$

Since R is positive definite and $(\phi-M_1y)^{\mathrm{T}}[(M_2-M_1)^{-1}(\phi-M_1y)-y] \leq 0$ it follows that $\dot{V}(x)$ is negative definite. $\square$

In order to specialize the result of Theorem 4.3 to robust stability with constant linear parameter uncertainty, consider the system

$$\dot{x}(t) = (A+\Delta A)x(t), \tag{4.23}$$

14

where $\Delta A \in \mathcal{U}$ and $\mathcal{U}$ is defined by

$$\mathcal{U} \triangleq \{\Delta A: \Delta A = -BFC, \quad F = \mathrm{diag}[F_1, F_2, \ldots, F_{m_0}], \quad \underline{m}_i \le F_i \le \overline{m}_i, \ i = 1, \ldots, m_0\}.$$

It now follows from Theorem 4.3 by setting $\phi(y) = Fy = FCx$ that $A + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$.

It has thus been shown that in the special case that F and N are diagonal nonnegative-definite matrices, Theorem 4.1 (with B_0 replaced by $-B_0$) specializes to the multivariable parabola criterion when applied to linear parameter uncertainty. This is not surprising since in this case the Lyapunov function (4.21) that establishes robust stability takes the form

$$V(x) = x^T P x + 2 \sum_{i=1}^{m_0} \int_0^{y_i} (F_i - \underline{m}_i)\sigma N_i d\sigma, \quad y_i = (C_0 x)_i, \tag{4.24}$$

or, equivalently,

$$V(x) = x^T P x + x^T C_0^T (F - M_1) N C_0 x \tag{4.25}$$

and thus is a special case of the parameter-dependent Lyapunov function discussed earlier. Note that the uncertain parameters are not allowed to be arbitrarily time-varying, which is consistent with the fact that the classical parabola criterion is restricted to *time-invariant* nonlinearities.

Finally, we note that in the case in which $M_1 = 0$, Theorem 4.3 specializes to the multivariable Popov criterion considered in [7,8]. Alternatively, retaining M_1 and setting $N = 0$ yields a strongly positive real requirement on $(M_2 - M_1)^{-1} + (I + G(s)M_1)^{-1}G(s)$ or, equivalently, $[(I + G(s)M_2)(I + G(s)M_1)^{-1}]$ which corresponds to the multivariable circle criterion considered in [6] with the restrictions that M_1, M_2 be diagonal and positive-definite.

5. Robust Controller Synthesis via the Parabola Riccati Equation

In this section we introduce the Robust Stability and Performance Problem with static output feedback control. This problem involves a set $\mathcal{U} \subset \mathbb{R}^{n \times n}$ of uncertain perturbations ΔA of the nominal system matrix A.

Robust Stability and Performance Problem. Given the nth-order stabilizable plant with constant real-valued plant parameter variations

$$\dot{x}(t) = (A + \Delta A)x(t) + Bu(t) + Dw(t), \ t \in [0, \infty), \tag{5.1}$$

$$y(t) = Cx(t), \tag{5.2}$$

15

where $u(t) \in \mathbb{R}^m, w(t) \in \mathbb{R}^d$, and $y(t) \in \mathbb{R}^\ell$, determine a static output feedback control law

$$u(t) = Ky(t) \tag{5.3}$$

that satisfies the following design criteria:

i) the closed-loop system (5.1) - (5.3) is asymptotically stable for all $\Delta A \in \mathcal{U}$, that is, $A + BKC + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$; and

ii) the performance functional

$$J(K) \triangleq \sup_{\Delta A \in \mathcal{U}} \limsup_{t \to \infty} \frac{1}{t} \mathbb{E} \left\{ \int_0^t [x^T(s)R_1 x(s) + u^T(s)R_2 u(s)] ds \right\} \tag{5.4}$$

is minimized.

For each variation $\Delta A \in \mathcal{U}$, the closed-loop system (5.1)-(5.3) can be written as

$$\dot{x}(t) = (\hat{A} + \Delta A)x(t) + Dw(t), \ t \in [0, \infty), \tag{5.5}$$

where

$$\hat{A} \triangleq A + BKC, \tag{5.6}$$

and where the white noise disturbance has intensity $V = DD^T$. Finally, note if $\hat{A} + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$ for a given K, then (5.4) can be written as

$$J(K) = \sup_{\Delta A \in \mathcal{U}} \text{tr } P_{\Delta A} V, \tag{5.7}$$

where $P_{\Delta A}$ satisfies (2.6) with A replaced by $\hat{A}$ and R replaced by

$$\hat{R} \triangleq R_1 + C^T K^T R_2 KC. \tag{5.8}$$

To apply Theorem 4.1 to controller synthesis we consider the performance bound (3.9) in place of the actual worst-case H$_2$ performance as in Theorem 4.1 with A, R replaced by $\hat{A}$ and $\hat{R}$ to address the closed-loop control problem. This leads to the following optimization problem.

Auxiliary Minimization Problem. Determine $K \in \mathbb{R}^{m \times l}$ that minimizes

$$\mathcal{J}(K) \triangleq \text{tr}[(P + C_0^T \mu C_0)V] \tag{5.9}$$

16

subject to

$$0 = (\hat{A} + B_0 M_1 C_0)^{\mathrm{T}} P + P(\hat{A} + B_0 M_1 C_0)$$
$$+ [C_0 + N C_0(\hat{A} + B_0 M_1 C_0) + B_0^{\mathrm{T}} P]^{\mathrm{T}} [(M_2 - M_1)^{-1} - N C_0 B_0 + ((M_2 - M_1)^{-1} - N C_0 B_0)^{\mathrm{T}}]^{-1}$$
$$\cdot [C_0 + N C_0(\hat{A} + B_0 M_1 C_0) + B_0^{\mathrm{T}} P] + \hat{R}. \tag{5.10}$$

It follows from Theorem 4.1 that the satisfaction of (5.10) along with the detectability condition $(\hat{A} + \Delta A, \hat{R})$ leads to closed-loop robust stability along with robust II$_2$ performance.

Next, we present sufficient conditions for robust stability and performance for the static output feedback case. For arbitrary $P, Q \in \mathbb{R}^{n \times n}$ define the notation

$$R_0 \triangleq (M_2 - M_1)^{-1} - N C_0 B_0 + ((M_2 - M_1)^{-1} - N C_0 B_0)^{\mathrm{T}},$$
$$R_{2a} \triangleq R_2 + B^{\mathrm{T}} C_0^{\mathrm{T}} N^{\mathrm{T}} R_0^{-1} N C_0 B,$$
$$P_a \triangleq B^{\mathrm{T}} P + B^{\mathrm{T}} C_0^{\mathrm{T}} N^{\mathrm{T}} R_0^{-1} C_0 + B^{\mathrm{T}} C_0^{\mathrm{T}} N^{\mathrm{T}} R_0^{-1} N C_0(A + B_0 M_1 C_0) + B^{\mathrm{T}} C_0^{\mathrm{T}} N^{\mathrm{T}} R_0^{-1} B_0^{\mathrm{T}} P,$$
$$\nu \triangleq Q C^{\mathrm{T}} (C Q C^{\mathrm{T}})^{-1} C, \qquad \nu_\perp \triangleq I_n - \nu,$$

when the indicated inverses exist.

Theorem 5.1. Assume $R_0 > 0$ and assume $\mathcal{F}$ and N are strongly compatible. Furthermore, suppose there exist $n \times n$ nonnegative-definite matrices P, Q such that $C Q C^{\mathrm{T}} > 0$ and

$$0 = [A + B_0 M_1 C_0 + B_0 R_0^{-1} C_0 + B_0 R_0^{-1} N C_0(A + B_0 M_1 C_0)]^{\mathrm{T}} P$$
$$+ P[A + B_0 M_1 C_0 + B_0 R_0^{-1} C_0 + B_0 R_0^{-1} N C_0(A + B_0 M_1 C_0)] + R_1$$
$$+ [C_0 + N C_0(A + B_0 M_1 C_0)]^{\mathrm{T}} R_0^{-1} [C_0 + N C_0(A + B_0 M_1 C_0)]$$
$$+ P B_0 R_0^{-1} B_0^{\mathrm{T}} P - P_a^{\mathrm{T}} R_{2a}^{-1} P_a - \nu_\perp^{\mathrm{T}} P_a^{\mathrm{T}} R_{2a}^{-1} P_a \nu_\perp, \tag{5.11}$$
$$0 = [A - B R_{2a}^{-1} P_a \nu + B_0 M_1 C_0 + B_0 R_0^{-1} N C_0(A - B R_{2a}^{-1} P_a \nu + B_0 M_1 C_0) + B_0 R_0^{-1} C_0 + B_0 R_0^{-1} B_0^{\mathrm{T}} P] Q$$
$$+ Q[A - B R_{2a}^{-1} P_a \nu + B_0 M_1 C_0 + B_0 R_0^{-1} N C_0(A - B R_{2a}^{-1} P_a \nu + B_0 M_1 C_0)$$
$$+ B_0 R_0^{-1} C_0 + B_0 R_0^{-1} B_0^{\mathrm{T}} P]^{\mathrm{T}} + V, \tag{5.12}$$

and let K be given by

$$K = -R_{2a}^{-1} P_a Q C^{\mathrm{T}} (C Q C^{\mathrm{T}})^{-1}. \tag{5.13}$$

Then $(\hat{A} + \Delta A, \hat{R})$ is detectable for all $\Delta A \in \mathcal{U}$ if and only if $\hat{A} + \Delta A$ is asymptotically stable for all $\Delta A \in \mathcal{U}$. In this case the closed-loop system performance (5.7) satisfies the bound

$$J(K) \leq \mathrm{tr}[(P + C_0^{\mathrm{T}} \mu C_0) V]. \tag{5.14}$$

Proof. The proof follows as in the proof given in [7]. □

Theorem 6.1 provides constructive sufficient conditions that yield static output feedback controllers for robust stability and performance. These conditions comprise a system of one modified algebraic Riccati equation and one modified Lyapunov equation in variables P and Q, respectively. Finally, note when solving (5.11) and (5.12) numerically, the matrices M_1, M_2 and N and the structure matrices B_0 and C_0 appearing in the design equations can be adjusted to examine tradeoffs between performance and robustness. To further reduce conservatism, one can view the multiplier matrix N as a free parameter and optimize N with respect to the worst case H_2 performance bound $\mathcal{J}$. In particular, computing $2J/2\mathcal{J} = 0$ yields

$$
\begin{aligned}
0 =&\, 1/2\mu C_0 V C_0^{\mathrm{T}} + [(M^{-1} - NC_0B_0) + (M^{-1} - NC_0B_0)]^{-1} \\
&\cdot [C_0 + NC_0(\hat{A} + B_0M_1C_0) + B_0^{\mathrm{T}}P]Q(\hat{A} + B_0M_1C_0)^{\mathrm{T}}C_0^{\mathrm{T}} \\
&+ [(M^{-1} - NC_0B_0) + (M^{-1} - NC_0B_0)^{\mathrm{T}}]^{-1}[C_0 + NC_0(\hat{A} + B_0M_1C_0) + B_0^{\mathrm{T}}P]Q \\
&\cdot [C_0 + NC_0(\hat{A} + B_0M_1C_0) + B_0^{\mathrm{T}}P]^{T}[(M^{-1} - NC_0B_0) + (M^{-1} - NC_0B_0)^{\mathrm{T}}]^{-1}B_0^{\mathrm{T}}C_0^{\mathrm{T}}.
\end{aligned}
$$

Now, the basic approach is to design the controller for a given N and then compute the optimal value of N for that controller. Hence, this design procedure will involve an interaction between controller design and evaluation of the multiplier N until convergence in N is achieved.

Next, we specialize Theorem 5.1 to the full-state feedback case. When the full state is available, that is, $C = I_n$, the projection $\nu = I_n$ so that $\nu_\perp = 0$. In this case (5.13) becomes

$$
K = -R_{2a}^{-1}P_a \tag{5.15}
$$

and (5.11), (5.12) collapse to the single equation

$$
\begin{aligned}
0 =&\, [A + B_0M_1C_0 + B_0R_0^{-1}C_0 + B_0R_0^{-1}NC_0(A + B_0M_1C_0)]^{\mathrm{T}}P \\
&+ P[A + B_0M_1C_0 + B_0R_0^{-1}C_0 + B_0R_0^{-1}NC_0(A + B_0M_1C_0)] + R_1 \\
&+ [C_0 + NC_0(A + B_0M_1C_0)]^{\mathrm{T}}R_0^{-1}[C_0 + NC_0(A + B_0M_1C_0)] \\
&+ PB_0R_0^{-1}B_0^{\mathrm{T}}P - P_a^{\mathrm{T}}R_{2a}^{-1}P_a. \tag{5.16}
\end{aligned}
$$

6. Dynamic Output Feedback Controller Synthesis

In this section we introduce the Dynamic Robust Stability and Performance Problem. For simplicity we restrict our attention to controllers of order $n_c = n$, that is, controllers whose order is equal to the dimension of the plant

Dynamic Robust Stability and Performance Problem. Given the nth-order stabilizable and detectable plant with constant structured real-valued plant parameter variations

$$\dot{x}(t) = (A + \Delta A)x(t) + Bu(t) + D_1 w(t), \ t \geq 0, \tag{6.1}$$

$$y(t) = Cx(t) + D_2 w(t), \tag{6.2}$$

where $u(t) \in \mathbb{R}^m, w(t) \in \mathbb{R}^d$, and $y(t) \in \mathbb{R}^\ell$, determine an nth-order dynamic compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \tag{6.3}$$

$$u(t) = C_c x_c(t), \tag{6.4}$$

that satisfies the following design criteria:

i) the closed-loop system (6.1)-(6.4) is asymptotically stable for all $\Delta A \in \mathcal{U}$; and

ii) the performance functional (5.4) with $J(K)$ replaced by $J(A_c, B_c, C_c)$ is minimized.

For each uncertain variation $\Delta A \in \mathcal{U}$, the closed-loop system (6.1)-(6.4) can be written as

$$\dot{\tilde{x}}(t) = (\tilde{A} + \Delta\tilde{A})\tilde{x}(t) + \tilde{D}w(t), \quad t \geq 0, \tag{6.5}$$

where

$$\tilde{x}(t) \triangleq \begin{bmatrix} x(t) \\ x_c(t) \end{bmatrix}, \ \tilde{A} \triangleq \begin{bmatrix} A & BC_c \\ B_c C & A_c \end{bmatrix}, \quad \Delta\tilde{A} = \begin{bmatrix} \Delta A & 0_{n \times n_c} \\ 0_{n_c \times n} & 0_{n_c \times n_c} \end{bmatrix},$$

and where the closed-loop disturbance $\tilde{D}w(t)$ has intensity

$$\tilde{V} = \tilde{D}\tilde{D}^{\mathrm{T}},$$

where $\tilde{D} \triangleq \begin{bmatrix} D_1 \\ B_c D_2 \end{bmatrix}, \tilde{V} \triangleq \begin{bmatrix} V_1 & 0 \\ 0 & B_c V_2 B_c^{\mathrm{T}} \end{bmatrix}, V_1 = D_1 D_1^{\mathrm{T}}, V_2 = D_2 D_2^{\mathrm{T}}$. The closed-loop system uncertainty $\Delta\tilde{A}$ has the form

$$\Delta\tilde{A} = \tilde{B}_0 F \tilde{C}_0 \tag{6.6}$$

where

$$\tilde{B}_0 \triangleq \begin{bmatrix} B_0 \\ 0_{n_c \times m_0} \end{bmatrix}, \quad \tilde{C}_0 \triangleq [C_0 \ \ 0_{m_0 \times n_c}].$$

19

Finally, if $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable for all $\Delta A \in \mathcal{U}$ for a given compensator (A_c, B_c, C_c), then the performance measure (5.4) is given by

$$J(A_c, B_c, C_c) = \sup_{\Delta A \in \mathcal{U}} \text{tr } \hat{P}_{\Delta\tilde{A}} \tilde{V}, \tag{6.7}$$

where $P_{\Delta\tilde{A}}$ satisfies the $2n \times 2n$ algebraic Lyapunov equation

$$0 = (\tilde{A} + \Delta\tilde{A})^\mathrm{T} \hat{P}_{\Delta\tilde{A}} + \hat{P}_{\Delta\tilde{A}}(\tilde{A} + \Delta\tilde{A}) + \tilde{R}, \tag{6.8}$$

where

$$\tilde{E} = [E_1 \quad E_2 C_c], \quad \tilde{R} = \tilde{E}^\mathrm{T}\tilde{E} = \begin{bmatrix} R_1 & 0 \\ 0 & C_c^\mathrm{T} R_2 C_c \end{bmatrix}.$$

Next, we proceed as in Section 5 where we replace the Lyapunov equation (6.8) for the dynamic problem with a Riccati equation that guarantees that the closed-loop system is robustly stable. Thus for the dynamic output feedback problem, Theorem 4.1 holds with A, R, V replaced by $\tilde{A}, \tilde{R}, \tilde{V}$.

For convenience in stating the main result of this section, recall the definitions of R_0, R_{2a}, P_a and define the additional notation

$$\bar{\Sigma} \triangleq C^\mathrm{T} V_2^{-1} C,$$

$$A_Q \triangleq A - Q\bar{\Sigma} + B_0 R_0^{-1} N C_0(A + B_0 M_1 C_0) + B_0 R_0^{-1} B_0^\mathrm{T} P + B_0 R_0^{-1} C_0 + B_0 M_1 C_0,$$

for arbitrary $Q, P \in \mathbb{R}^{n \times n}$.

Theorem 6.1. Assume $R_0 > 0$ and assume $\mathcal{F}$ and N are strongly compatible. Furthermore, suppose there exist $n \times n$ nonnegative-definite matrices $P, Q, \hat{P}$ satisfying

$$\begin{aligned} 0 =& [A + B_0 M_1 C_0 + B_0 R_0^{-1} C_0 + B_0 R_0^{-1} N C_0(A + B_0 M_1 C_0)]^\mathrm{T} P \\ &+ P[A + B_0 M_1 C_0 + B_0 R_0^{-1} C_0 + B_0 R_0^{-1} N C_0(A + B_0 M_1 C_0)] + R_1 \\ &+ [C_0 + N C_0(A + B_0 M_1 C_0)]^\mathrm{T} R_0^{-1}[C_0 + N C_0(A + B_0 M_1 C_0)] \\ &+ P B_0 R_0^{-1} B_0^\mathrm{T} P - P_a^\mathrm{T} R_{2a}^{-1} P_a, \end{aligned} \tag{6.9}$$

$$0 = [A + B_0 M_1 C_0 + B_0 R_0^{-1} B_0^\mathrm{T}(P + \hat{P}) + B_0 R_0^{-1} N C_0(A + B_0 M_1 C_0)]Q \tag{6.10}$$
$$+ Q[A + B_0 M_1 C_0 + B_0 R_0^{-1} B_0^\mathrm{T}(P + \hat{P}) + B_0 R_0^{-1} N C_0(A + B_0 M_1 C_0)]^\mathrm{T} + V_1 - Q\bar{\Sigma}Q,$$

$$0 = A_Q^\mathrm{T} \hat{P} + \hat{P} A_Q + \hat{P} B_0 R_0^{-1} B_0^\mathrm{T} \hat{P} + P_a^\mathrm{T} R_{2a}^{-1} P_a, \tag{6.11}$$

and let A_c, B_c, C_c be given by

$$A_c = A - Q\bar{\Sigma} - BR_{2a}^{-1}P_a - B_0 R_0^{-1} N C_0 B R_{2a}^{-1} P_a$$
$$+ B_0 R_0^{-1} N C_0 (A + B_0 M_1 C_0) + B_0 R_0^{-1} C_0 + B_0 M_1 C_0 + B_0 R_0^{-1} B_0^T P, \qquad (6.12)$$

$$B_c = Q C^T V_2^{-1}, \qquad (6.13)$$

$$C_c = - R_{2a}^{-1} P_a. \qquad (6.14)$$

Then $(\tilde{A} + \Delta\tilde{A}, \tilde{E})$ is detectable for all $\Delta A \in \mathcal{U}$ if and only if $\tilde{A} + \Delta\tilde{A}$ is asymptotically stable for all $\Delta A \in \mathcal{U}$. In this case, the performance of the closed-loop system (6.5) satisfied the bound

$$J(A_c, B_c, C_c) \leq \mathrm{tr}[(P + \hat{P})V_1 + \hat{P}Q\bar{\Sigma}Q + C_0^T \mu C_0 V_1] \qquad (6.15)$$

Proof. The proof follows as in the proof given in [7]. $\qquad \qquad \square$

Remark 6.1 Note that if the uncertainty in the plant dynamics is deleted, that is, $B_0 = 0$, $C_0 = 0$, then Theorem 6.1 specializes to the standard LQG result.

References

[1] B. D. O. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis: A Modern Systems Theory Approach,* Prentice-Hall, 1973.

[2] A. R. Bergen and M. A. Sapiro, "The parabola test for absolute stability," *IEEE Trans. Autom. Contr.,* Vol. AC-12, pp. 312-314, 1967.

[3] D. S. Bernstein, W. M. Haddad, and D. C. Hyland, "Small gain versus positive real modeling of real parameter uncertainty," *AIAA J. Guid. Contr. Dyn.,* Vol. 15, pp. 538-540, 1992.

[4] D. S. Bernstein, W. M. Haddad, D. C. Hyland, and F. Tyan, "A maximum entropy-type Lyapunov function for robust stability and performance analysis," *Syst. Contr. Lett.,* submitted.

[5] W. M. Haddad and D. S. Bernstein, "Robust stabilization with positive real uncertainty: Beyond the small gain theorem," *Syst. Contr. Lett.,* Vol. 17, pp. 191-208, 1991.

[6] W. M. Haddad and D. S. Bernstein, "Explicit construction of quadratic Lyapunov functions for the small gain, positivity, circle, and Popov theorems and their application to robust stability," in *Control of Uncertain Dynamic Systems,* S.P. Bhattacharyya and L. H. Keel, Eds., CRC Press, pp. 1249-173, 1991.

[7] W. M. Haddad and D. S. Bernstein, "Parameter-dependent Lyapunov functions, constant real parameter uncertainty, and the Popov criterion in robust analysis and synthesis Part 1, Part 2," *Proc. IEEE Conf. Dec. Contr.,* pp. 2274-2279, 2632-2633, Brighton, U.K., December 1991.

[8] W. M. Haddad and D. S. Bernstein, "Parameter-dependent Lyapunov functions, constant real parameter uncertainty, and the Popov criterion in robust analysis and synthesis," *IEEE Trans. Autom. Contr.,* submitted.

[9] R. Lozano-Leal and S. Joshi, "Strictly positive real transfer functions revisited," *IEEE Trans. Autom. Contr.,* Vol. 35, pp. 1243-1245, 1990.

[10] K. S. Narendra and J. H. Taylor, *Frequency Domain Criteria for Absolute Stability,* Academic Press, New York, 1973.

[11] J. T. Wen, "Time domain and frequency domain conditions for strict positive realness," *IEEE Trans. Autom. Contr.,* Vol. 33, pp. 988-992, 1988.

[12] W. M. Wonham, *Linear Multivariable Control: A Geometric Approach,* Springer-Verlag, New York, 1979.

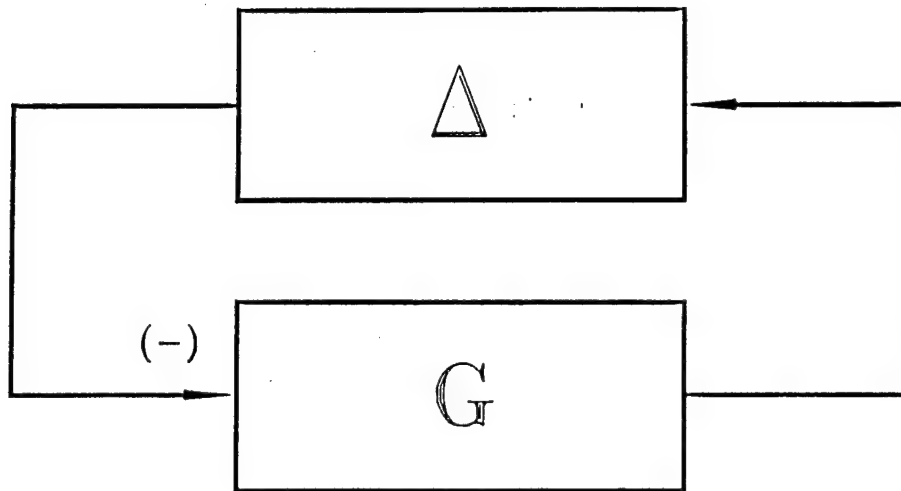Figure 1
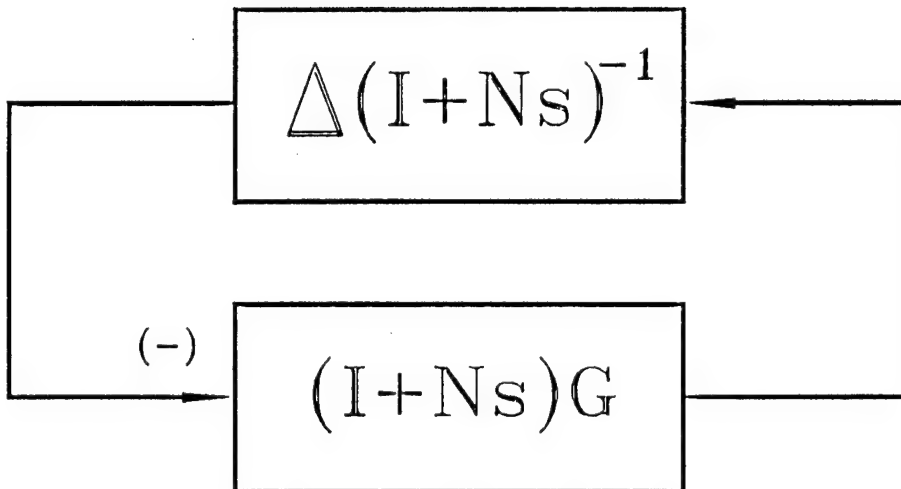

Figure 2
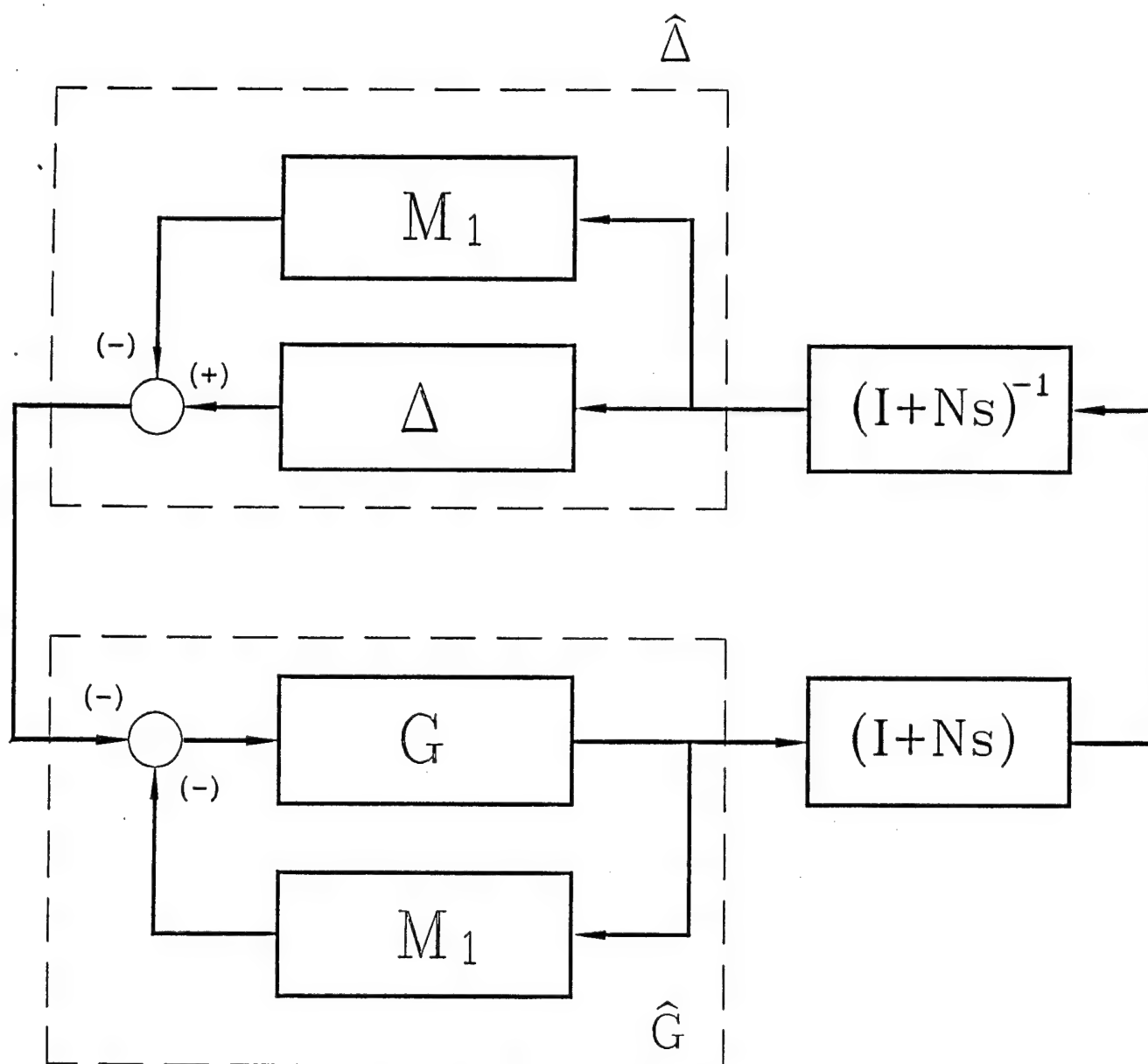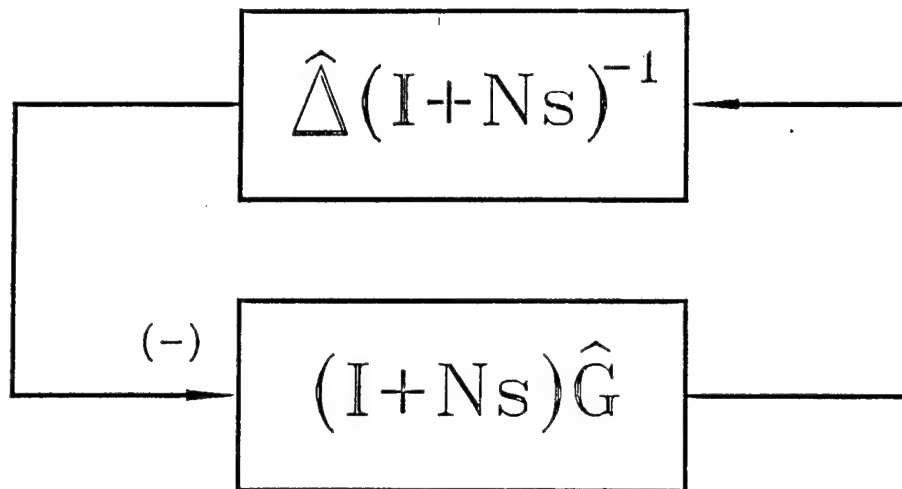
Figure 3

$$\hat{G} \;=\; (I+GM_1)^{-1}G, \qquad \hat{\Delta} \;=\; \Delta - M_1.$$

Figure 4

Appendix D:

Robust Stability Analysis Using the Small Gain, Circle Positivity, and Popov Theorems: A Comparative Study

Letters

Robust Stability Analysis Using the Small Gain, Circle, Positivity, and Popov Theorems: A Comparative Study

Wassim M. Haddad, Emmanuel G.
Collins, Jr., and Dennis S. Bernstein

Abstract—This note analyzes the stability robustness of a Maximum Entropy controller designed for a benchmark problem. Four robustness tests are used: small gain analysis, circle analysis, positive real analysis, and Popov analysis, each of which is guaranteed to give a less conservative result than the previous test. The analysis here is performed graphically although recent research has developed equivalent tests based on Lyapunov theory. The Popov test is seen, for this example, to yield highly nonconservative robust stability bounds. The results here illuminate the conservatism of analysis based on traditional small-gain type tests and reveal the effectiveness of analysis tests based on Popov analysis and related parameter-dependent Lyapunov functions.

I. INTRODUCTION

In control engineering practice, control design (whether classical or modern) is usually predicated upon some nominal (usually linear) model of the plant to be controlled. However, this nominal model of the system is never an exact representation of the true physical system. This necessitates tools that allow a control system to be analyzed for robustness with respect to errors in the design model. These analysis tools almost always lead to techniques for actually designing a control system for robustness.

In classical control, gain and phase margins are often used as indirect measures of robustness. However, these criteria do not always adequately provide robustness with respect to the true plant uncertainties. Hence, to add reliability to the analysis process, more direct and rigorous measures of robustness are needed. To guarantee the best performance possible, in the presence of uncertainties in the system model, it is important that these robustness measures be nonconservative.

In the analysis of systems for robustness, the conservatism of the resulting robust stability and performance bounds is largely dependent upon the characterization of the uncertainty in the analysis process. This uncertainty characterization can be viewed as lying between two extremes. In the state space, one extreme would be to model the uncertainty as constant, real parameters while the opposite extreme would be to model the uncertainty as arbitrarily time-varying, real parameters. In the frequency domain, the corresponding extremes are to model the uncertainty as a transfer function with bounded phase or oppositely, as a transfer function with arbitrary phase.

If the uncertainty is truly constant and real, then modeling it as arbitrarily time-varying can lead to very conservative results. For example, classical analysis of a Hill's equation (e.g., the Mathieu

Wassim M. Haddad is with the Department of Mechanical and Aerospace Engineering, Florida Institute of Technology, Melbourne, FL 32901.

Emmanuel G. Collins, Jr. is with the Harris Corporation, Government Aerospace Systems, MS 1914849 Melbourne, FL 32902.
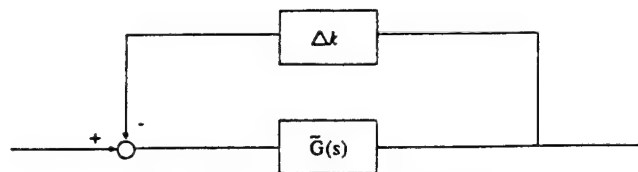
Fig. 1 Standard uncertainty representation.

equation) shows that time-varying parameter variations can destabilize a system even when the parameter variations are confined to a region in which constant variations are nondestabilizing (1). Also, as seen in [2] which analyzes stiffness uncertainty for a flexible structure, when uncertainty is modeled as having arbitrary phase, predictions for stability and performance will be much more conservative than results developed assuming phase-bounded (e.g., positive real) uncertainty.

In recent years it has become conventional to model plant uncertainty, say Δk, using the feedback configuration shown in Figure 1. In this figure $\hat{G}(s)$ denotes the nominal plant. Four of the most fundamental results concerning stability of feedback system interconnections are the small gain, circle, positivity, and Popov theorems [1, 3]. Even though these theorems were originally developed to analyze stability of system with a single, memoryless nonlinear element in a feedback configuration [1], in recent research [3, 4] each result was reinterpreted and generalized to the problem of robust stability involving linear uncertainty. To do this, a Lyapunov function framework was established, providing connections of these classical results to robust stability and performance via state space methods.

As shown in [3], the main difference between the small gain, circle, and positivity theorems versus the Popov theorem is that the former results guarantee robustness with respect to arbitrarily, time-varying uncertainty while the Popov theorem restricts the time variation of the uncertainty. This is not surprising once one recognizes that the Lyapunov function foundation of the small gain, circle, and positivity theorems is based upon conventional or "fixed" quadratic Lyapunov functions which, of course, guarantee stability with respect to arbitrarily, time-varying perturbations. In contrast, the Popov theorem is based upon a quadratic Lyapunov function that is a function of the parametric uncertainty, that is, a parameter-dependent quadratic Lyapunov function [3, 4]. Hence, in effect, the Popov result guarantees stability by means of a *family* of Lyapunov functions. A key aspect of this approach [4] is the fact that it does *not* apply to arbitrarily time-varying uncertainties, which renders it significantly less conservative than fixed quadratic Lyapunov functions in the presence of constant real parameter uncertainty.

To illuminate the conservatism of robustness analysis based on traditional small-gain type tests for constant real parameter uncertainty and to reveal the importance of tests which restrict the time-variation in the state space and thus allow the incorporation of phase information in the frequency domain, we consider a simple two-mass/spring, lightly damped, system with uncertain stiffness [5]. This example was chosen to highlight the inherent drawbacks of small gain principles applied to the analysis of feedback systems with constant real parameter uncertainty. A quadratic Lyapunov function framework leading to an algebraic basis in terms of matrix Riccati

Fig. 2 Spring-mass system.

equations for the analysis and synthesis of robust controllers for the small-gain, circle, positivity, and Popov theorems is given in [3, 4]. Nevertheless, for simplicity the analysis presented here is graphical.

II. TWO-MASS/SPRING EXAMPLE

Consider the two-mass/spring system shown in Figure 2 with uncertain stiffness k. A control force acts on body 1, and the position of body 2 is measured resulting in a noncolocated control problem. Here, we consider Controller \#1 of [6, 7] which was designed for Problem \# 1 of a benchmark problem [5] using the Maximum Entropy robust control design technique. The controller transfer function given by

$$H(s) = \frac{194390(s + 0.33679)[(s - 0.11735)^2 + 0.90996^2]}{(s + 81.438)(s + 131.04)[(s + 2.9049)^2 + 1.8615^2]} \quad (1)$$
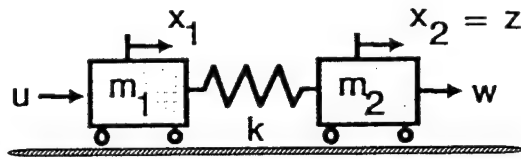
was designed so that the closed-loop system is robust with respect to perturbations in the nominal value of the stiffness k (i.e., $k = k_{nom}$). The exact stiffness stability region over which the system will remain stable was computed by a simple search and is given by

$$0.4459 \le k \le 2.0660. \quad (2)$$

Next, using a graphical approach we apply small gain analysis, circle analysis, positive real analysis, and Popov analysis to determine the stiffness stability regions predicted by each of these tests. Each of these tests is related to the previous test and is guaranteed to be less conservative.

We begin by constructing the uncertainty feedback system that will be used in each of the tests. The plant (for $m_1 = m_2 = 1$) is given by the triple $(A(k), B, C)$ where

$$A(k) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k & k & 0 & 0 \\ k & -k & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \quad C = [0\ 1\ 0\ 0]. \quad (3)$$

The perturbation in $A(k)$ due to a change in the stiffness element k from nominal value k_{nom} is given by

$$A(k) - A(k_{nom}) \triangleq \Delta A = B_0 \Delta k C_0 \quad (4)$$

where $B_0^{T} = [0\ 0\ -1\ 1]$ and $C_0 = [1\ -1\ 0\ 0]$. In the subsequent analysis we will choose $k_{nom} = 0.6$ since the controller (1) was developed under this assumption.

Let the triple (A_c, B_c, C_c) denote the state space representation of the controller (1). Then, assuming negative feedback, the closed-loop state matrix is given by

$$\tilde{A}(k) = \begin{bmatrix} A(k) & -BC_c \\ BC_c & A_c \end{bmatrix}. \quad (5)$$
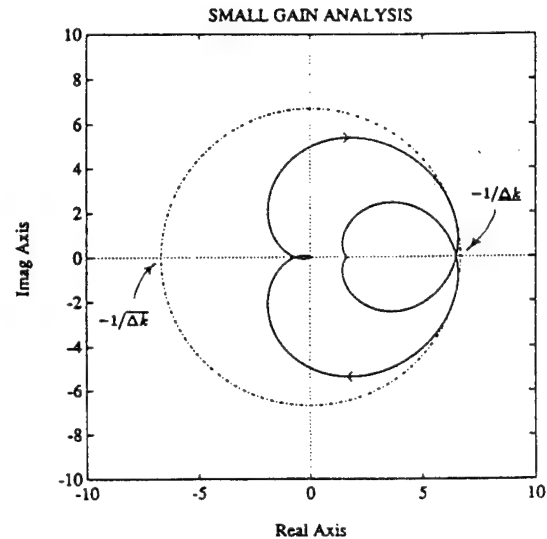
Fig. 3 Small gain analysis.

Next, define $\tilde{B}_0^{T} = [B_0^{T}\ 0_{1\times4}]$, $\tilde{C}_0 = [C_0\ 0_{1\times4}]$ and let $\tilde{G}(s) = -\tilde{C}_0(sI - \tilde{A}(k_{nom}))^{-1}\tilde{B}_0$. Then, the plant uncertainty Δk can be represented by a fictitious feedback loop as shown in Figure 1.

For each of the tests below we will determine $\overline{\Delta k}$ (positive) and $\underline{\Delta k}$ (negative) such that stability is guaranteed for

$$k_{nom} + \underline{\Delta k} \le k \le k_{nom} + \overline{\Delta k}. \quad (6)$$

Small Gain Analysis

Small gain analysis requires considering the Nyquist diagram of $\tilde{G}(s)$. The smallest circle centered at the origin that completely encompasses the Nyquist diagram, $Im[\tilde{G}(j\omega)]$ vs. $Re[\tilde{G}(j\omega)]$ for all ω, (without touching it) is then drawn. The intersection of this circle with the negative real axis is given by $-1/\overline{\Delta k}$ and the intersection with the positive real axis is given by $-1/\underline{\Delta k}$. This analysis is shown in Figure 3. It follows that $\overline{\Delta k} = 0.1496$ and $\underline{\Delta k} = -0.1496$. Hence, using small gain analysis, stability is guaranteed for

$$0.4504 \le k \le 0.7496. \quad (7)$$

Note that since the Δk uncertainty block is comprised of a single scalar, this result is equivalent to a μ-analysis test [8].

Circle Analysis

As in small gain analysis, circle analysis determines stability bounds by drawing a circle that completely encompasses the Nyquist diagram (without touching it). However, the circle criterion allows the center of the circle to lie anywhere along the real axis and can hence give a less conservative bound $\overline{\Delta k}$ (or $\underline{\Delta k}$) at the expense of increased conservatism in the remaining bound $\underline{\Delta k}$ (or $\overline{\Delta k}$). Here we choose the center of the circle to lie at $((x_{min} + x_{max})/2, 0)$ where x_{min} is the minimum real part of the Nyquist diagram and x_{max} is the maximum real part. The intersection of this circle with the negative real axis equals $-1/\overline{\Delta k}$ and the intersection with the positive real axis equals by $-1/\underline{\Delta k}$. This analysis is shown in Figure 4. It follows that $\overline{\Delta k} = 0.3167$ and $\underline{\Delta k} = -0.1277$. Hence, using circle analysis, stability is guaranteed for

$$0.4722 \le k \le 0.9167. \quad (8)$$

Fig. 4 Circle analysis.

Fig. 5 Positive real analysis.

Fig. 6 Popov analysis.
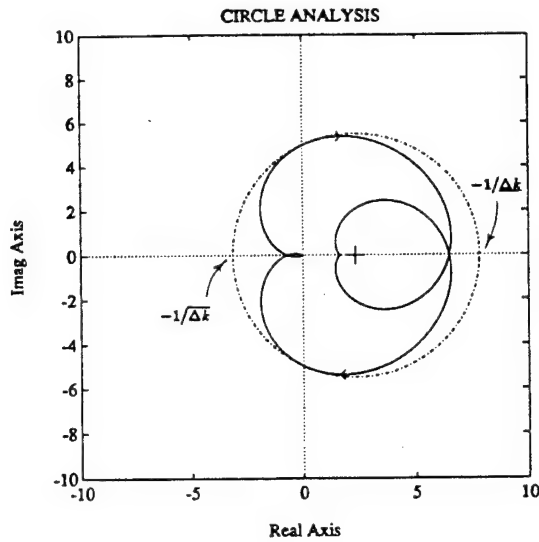
Positive Real Analysis

Positive real analysis determines stability bounds by drawing straight-lines that lie to the left or right of the Nyquist diagram (without touching it). It is equivalent to the limit of the circle criterion as the center of the circle moves toward infinity along the positive or negative real axis and will always give less conservative bounds. For the Nyquist diagram of $\hat{G}(s)$, the intersection of the line to the left of the Nyquist plot with the negative real axis equals $-1/\overline{\Delta k}$. The intersection of the line to the right of the Nyquist plot with the positive real axis equals $-1/\underline{\Delta k}$. This analysis is shown in Figure 5. It follows that $\overline{\Delta k} = 0.5277$ and $\underline{\Delta k} = -0.1522$. Hence, using positive real analysis, stability is guaranteed for

$$0.4478 \leq k \leq 1.1277. \tag{9}$$

Popov Analysis

Popov analysis is a test that determines a stability bound from a modified Nyquist diagram, namely the Popov plot, $\omega Im[\hat{G}(j\omega)]$ vs. $Re[\hat{G}(j\omega)]$ for $\omega \geq 0$. This analysis requires finding lines (Popov lines) that intersect the negative or positive real axis at a point that is to the left of the Popov plot but as close to the origin as possible. The slope of these lines are $-1/\overline{N}$ and $-1/\underline{N}$ where $\overline{N}$ and $\underline{N}$ are the Popov multipliers. The Popov test is equivalent to the positive real test if the lines are chosen to be vertical. For the Popov diagram of $\hat{G}(s)$, the intersection of the line to the left of the Popov plot with the negative real axis equals $-1/\overline{\Delta k}$. The intersection of the line to the right of the Popov plot with the positive real axis equals $-1/\underline{\Delta k}$. This analysis is shown in Figure 6. It follows that $\overline{\Delta k} = 1.4660$ and $\underline{\Delta k} = 0.1541$ and the corresponding Popov multipliers are respectively $\overline{N} = 0.7999$ and $\underline{N} = -0.2755$. Hence, using Popov analysis, stability is guaranteed for

$$0.4459 \leq k \leq 2.0660. \tag{10}$$

Note that these bounds are identical to the exact bounds (2), at least to four-digit precision for the lower bound and five digit precision for the upper bound. Hence, for this example, Popov
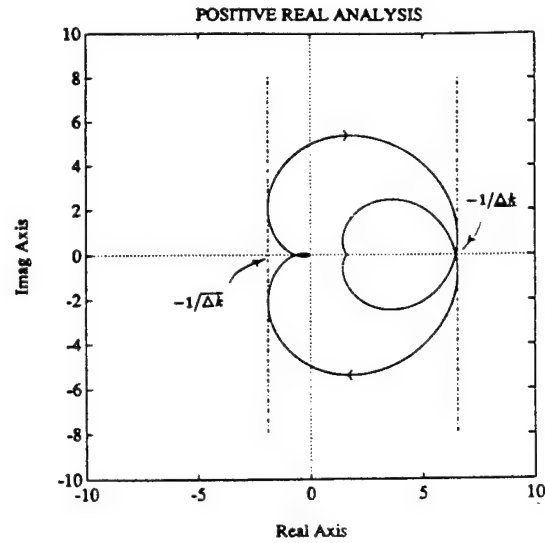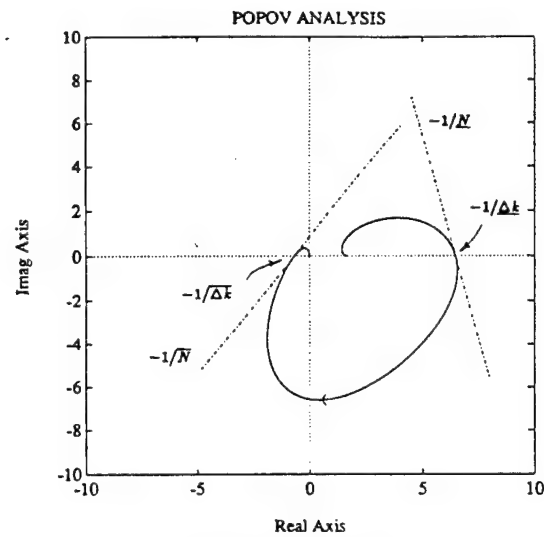
analysis yielded highly nonconservative results. This is not surprising since, as mentioned in the Introduction, the Popov result is based upon a parameter-dependent Lyapunov function which severely restricts the allowable time variation of the uncertain parameters and hence closely approximates real parameter uncertainty within robustness analysis.

III. CONCLUSION

We have shown by means of a simple two-mass/spring example with uncertain stiffness that small gain modeling of constant real parameter uncertainty can be extremely conservative. An alternative approach to the phase information/real parameter uncertainty problem using Popov analysis and related parameter-dependent Lyapunov functions was shown to be significantly less conservative. Although Popov analysis was traditionally developed to analyze stability of a system with a single, memoryless *nonlinear* element in a feedback configuration, recent results have reinterpreted Popov analysis to handle the problem of robust stability involving *constant, linear* uncertainty [3, 4].

The results here demonstrate the somewhat overlooked fact that Popov analysis can be very nonconservative when applied to the analysis of linear systems with linear uncertainty.

Finally, it should be acknowledged that the results of [3, 4] allow Popov analysis to be used to synthesize robust controllers. This problem of robust control can, of course, be alternatively approached using adaptive control techniques [9, 10] which implicitly or explicitly identify the model uncertainty. It is possible that the results discussed in [3, 4] can be used as a basis for using Popov analysis to determine the stability and robustness properties of adaptive controllers.

REFERENCES

[1] K. S. Narendra and J. H. Taylor, *Frequency Domain Criteria for Absolute Stability*, New York: Academic Press, 1973.

[2] D. S. Bernstein, W. M. Haddad, and D. C. Hyland "Small Gain Versus Positive Real Modeling of Real Parameter Uncertainty," *AIAA J. Guid. Contr. Dyn.*, vol. 15, pp. 538–540, 1992.

[3] W. M. Haddad and D. S. Bernstein, "Explicit Construction of Quadratic Lyapunov Functions for the Small Gain, Positivity, Circle, and Popov Theorems and Their Application to Robust Stability," *Proc. IEEE Conf. Dec. Contr.*, pp. 2618–2623, Brighton, U.K., December 1991. (Also submitted to *Int. J. Robust and Nonlinear Control*)

[4] W. M. Haddad and D. S. Bernstein, "Parameter-Dependent Lyapunov Functions, Constant Real Parameter Uncertainty, and the Popov Criterion in Robust Analysis and Synthesis: Part I, Part II," *Proc. IEEE Conf. Dec. Contr.*, Brighton, U.K., December 1991, pp. 2274–2279, pp. 2632–2633. (Also to be published in *IEEE Trans. Autom. Contr.*)

[5] B. Wie and D. S. Bernstein, "A Benchmark Problem for Robust Control Design," *Proc. Amer. Contr. Conf.*, San Diego, CA, May, 1990, pp. 961–962.

[6] E. G. Collins, Jr., J. A. King, and D. S. Bernstein, "Robust Control Design for a Benchmark Problem Using the Maximum Entropy Approach," *Proc. Amer. Contr. Conf.*, Boston, MA, June 1991, pp. 1935–1936.

[7] E. G. Collins, Jr., J. A. King, and D. S. Bernstein, "Application of Maximum Entropy/Optimal Projection Design Synthesis to the Benchmark Problem," *AIAA J. Guid. Contr. Dyn.*, to be published.

[8] J. C. Doyle, "Analysis of Feedback Systems with Structured Uncertainties," *IEE Proc.*, Part D, Vol. 129, 1982, pp. 242–250.

[9] K. J. Astrom and B. Wittenmark, *Adaptive Control*, New York: Addison Wesley, 1989.

[10] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems*, Englewood Cliffs, New Jersey: Prentice Hall, 1989.

Appendix E:

Riccati Equation Approaches for Robust Stability and Performance Analysis Using the Small Gain, Positivity, and Popov Theorems

Riccati Equation Approaches for Small Gain, Positivity, and Popov Robustness Analysis

Emmanuel G. Collins Jr.*
Harris Corporation, Melbourne, Florida 32902
Wassim M. Haddad†
Florida Institute of Technology, Melbourne, Florida 32901
and
Lawrence D. Davis*
Harris Corporation, Melbourne, Florida 32902

In recent years, small gain (or H_∞) analysis has been used to analyze feedback systems for robust stability and performance. However, since small gain analysis allows uncertainty with arbitrary phase in the frequency domain and arbitrary time variations in the time domain, it can be overly conservative for constant real parametric uncertainty. More recent results have led to the development of robustness analysis tools, such as extensions of Popov analysis, that are less conservative. These tests are based on parameter-dependent Lyapunov functions, in contrast to the small gain test, which is based on a fixed quadratic Lyapunov function. This paper uses a benchmark problem to compare Popov analysis with small gain analysis and positivity analysis (a special case of Popov analysis that corresponds to a fixed quadratic Lyapunov function). The state-space versions of these tests, based on Riccati equations, are implemented using continuation algorithms. The results show that the Popov test is significantly less conservative than the other two tests and for this example is completely nonconservative in terms of its prediction of robust stability.

I. Introduction

ONE of the most important aspects of the control design and evaluation process is the analysis of feedback systems for robust stability and performance. Over the past several years, significant attention has been devoted to the use of small gain (or H_∞) tests for robustness analysis.[1-5] However, it is well known that these tests can be very conservative since in the frequency domain the small gain test characterizes uncertainty with bounded gain but arbitrary phase, whereas in the time domain the small gain test characterizes uncertainty with arbitrary time variation.[5] This conservatism has led to the search for more accurate robustness tests. In particular, researchers have searched for tests that allow frequency domain uncertainty characterization to include phase bounding or time domain uncertainty characterization to include restrictions on the allowable time variations.

The small gain test is actually based on conventional or "fixed" quadratic Lyapunov functions that guarantee stability with respect to arbitrarily time-varying perturbations. Very recently, however, robustness tests have been developed that are based on quadratic Lyapunov functions that are a function of the parametric uncertainty, that is, "parameter-dependent Lyapunov functions."[6,7] In contrast to analysis based on a fixed quadratic Lyapunov function, these tests guarantee robust stability by means of a family of Lyapunov functions and do not apply to arbitrarily time-varying uncertainties. Hence, when the actual uncertainty is real and constant, these tests are less conservative than tests based on fixed quadratic Lyapunov functions.[6]

In this paper we use a benchmark problem to compare the Popov test,[2] based on a parameter-dependent Lyapunov function,[6,7] with the small gain[2] and positivity tests[2] that are based on fixed quadratic Lyapunov functions.[6,7] Each of the stability tests has graphical interpretations for the case of one-block, scalar uncertainty.[2] However, here we will emphasize the state-space tests that are based on Riccati equations and allow the development of robust H_2 performance bounds in addition to the determination of robust stability. We develop continuation algorithms for the special case of one-block, scalar uncertainty. The algorithm for Popov analysis additionally requires that a certain product ($\bar{C}_0 \bar{B}_0$) related to the uncertainty characterization be equal to zero. As will be seen in Sec. III, this condition holds for the parametric uncertainty under consideration. The algorithms are applied to analyze a feedback system for the benchmark system in which the controller was designed using the maximum entropy approach.[8]

The paper is organized as follows. Section II presents the linear system to be analyzed for robust stability and performance and gives the main theorems for the small gain, positivity, and Popov tests. Section III then considers the benchmark problem and formulates the feedback system to be analyzed in the format of Sec. II. Section IV applies the graphical tests to determine robust stability. Next, Sec. V develops continuation algorithms for a special case of the state-space tests and applies the algorithms to the benchmark problem. Finally, Sec. VI discusses the conclusions and directions for future work.

II. Riccati Equation Characterizations for the Small Gain, Positivity, and Popov Theorems

We begin this section by establishing some basic notation and definitions. Let $\Re$ denote the real numbers, and let $(\cdot)^T$ and $(\cdot)^*$ denote transpose and complex conjugate transpose. Furthermore, we write $\|\cdot\|_2$ for the Euclidean norm, $\|\cdot\|_F$ for the Frobenius norm, $\sigma_{max}(\cdot)$ for the maximum singular value, $\text{tr}(\cdot)$ for the trace operator, and $M \geq 0$ ($M > 0$) to denote the fact that the hermitian matrix M is nonnegative (positive) definite. The notation

$$G(s) \sim \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array} \right] \tag{1}$$

denotes that $G(s)$ is a transfer function corresponding to the state-space realization (A, B, C, D), i.e., $G(s) = C(sI - A)^{-1} \times B + D$. The notation "$\min$" is used to denote a minimal realization. For asymptotically stable $G(s)$, define the H_2 and H_∞ norms, respectively, where $\omega \in [0, \infty)$, as

$$\|G(s)\|_2^2 \triangleq \frac{1}{2\pi} \int_{-\infty}^{\infty} \|G(j\omega)\|_F^2 \, d\omega \tag{2a}$$

$$\|G(s)\|_\infty \triangleq \sigma_{max}(G(j\omega)) \tag{2b}$$

A transfer function $G(s)$ is *bounded real* if 1) $G(s)$ is asymptotically stable and 2) $\|G(j\omega)\|_\infty \leq 1$ for $\omega \in [0,\infty)$. Furthermore, $G(s)$ is called *strictly bounded real* if 1) $G(s)$ is asymptotically stable and 2) $\|G(j\omega)\|_\infty < 1$ for $\omega \in [0,\infty)$. Finally, note that if $G(s)$ is strictly bounded real, then $I - D^T D > 0$, where $D \triangleq G(\infty)$.

A square transfer function $G(s)$ is called *positive real* if 1) all poles of $G(s)$ are in the closed left half-plane and 2) $G(j\omega) + G^*(j\omega)$ is nonnegative definite for $\omega \in [0,\infty)$. A square transfer function $G(s)$ is called strictly positive real if 1) $G(s)$ is asymptotically stable and 2) $G(j\omega) + G^*(j\omega)$ is positive definite for $\omega \in [0,\infty)$. Finally, a square transfer function $G(s)$ is strongly positive real if it is *strictly positive real* and $D + D^T > 0$, where $D \triangleq G(\infty)$. (Note that in some of the literature "strongly positive real" as defined here is referred to as "strictly positive real.")

At this point, we consider a linear uncertain system of the form

$$\dot{\bar{x}}(t) = (\bar{A} - \bar{B}_0 F \bar{C}_0)\bar{x}(t) + \bar{D}w(t), \qquad \bar{x}(t) \in \Re^{\bar{n}} \tag{3}$$

$$z(t) = \bar{E}\bar{x}(t) \tag{4}$$

Note that the system (3) and (4) may denote a linear feedback system subject to an exogenous disturbance signal $w(t)$. The individual elements of $z(t)$ may denote the performance variables, possibly including the actuation signals. The product $-\bar{B}_0 F \bar{C}_0$ then denotes the parametric uncertainty (i.e., $\Delta\bar{A}$). In particular, $\bar{B}_0$ and $\bar{C}_0$ are fixed matrices denoting the structure of the uncertainty and F is an uncertain matrix. Here, it is assumed that for some nonnegative definite diagonal matrix M, $F \in F_M^+$, or for some nonnegative scalar γ, $F \in F_\gamma$ where

$$F_M^+ \triangleq \{F \in \Re^{m_0 \times m_0}: F \text{ is diagonal}, 0 \leq F \leq M\} \tag{5}$$

$$F_\gamma \triangleq \{F \in \Re^{m_0 \times m_0}: F \text{ is diagonal}, F^2 \leq \gamma^{-2} I_{m_0}\} \tag{6}$$

If we additionally define

$$F_M^- \triangleq \{F \in \Re^{m_0 \times m_0}: F \text{ is diagonal}, -M \leq F \leq 0\} \tag{7}$$

then $F \in F_M^+$ if and only if $-F \in F_M^-$, and if $\gamma^{-1} = \sigma_{max}(M)$, $F \in F_\gamma$ implies $F \in F_M^+ \cup F_M^-$.

Now, denote $\bar{G}(s)$ by

$$\bar{G}(s) \sim \left[\begin{array}{c|c} \bar{A} & \bar{B}_0 \\ \hline \bar{C}_0 & 0 \end{array} \right] \tag{8}$$

Then evaluation of the robust stability of Eq. (3) is equivalent to evaluation of the robust stability of the feedback system shown in Fig. 1.

It now follows that for asymptotically stable $\bar{A} - \bar{B}_0 F \bar{C}_0$ the H_2 norm for Eqs. (3) and (4) is given by

$$J(F) = \text{tr } \bar{Q}\bar{R} = \text{tr } \bar{P}\bar{V} \tag{9}$$

where

$$\bar{R} = \bar{E}^T \bar{E} \tag{10}$$

$$\bar{V} = \bar{D}\bar{D}^T \tag{11}$$
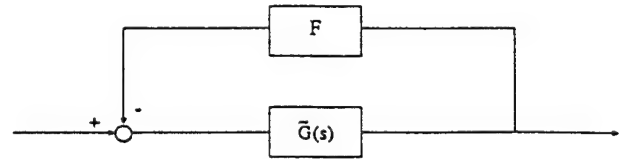
Fig. 1 Feedback system to be analyzed for robust stability.

$$0 = (\bar{A} - \bar{B}_0 F \bar{C}_0)^T \bar{P} + \bar{P}(\bar{A} - \bar{B}_0 F \bar{C}_0) + \bar{R} \tag{12}$$

$$0 = (\bar{A} - \bar{B}_0 F \bar{C}_0)\bar{Q} + \bar{Q}(\bar{A} - \bar{B}_0 F \bar{C}_0)^T + \bar{V} \tag{13}$$

If $w(t)$ is a standard white noise process with identity intensity, then $J(F) = \lim_{t \to \infty} \mathcal{E}[\bar{x}^T(t)\bar{R}\bar{x}(t)]$. Later we will present robust performance bounds $\bar{J}$ such that $J(F) \leq \bar{J}$ for each F in the uncertainty set.

Next, we state the versions of the small gain, positivity, and Popov theorems that give sufficient conditions for the stability of the uncertain system (3) or, equivalently, the negative feedback interconnection of Fig. 1. Each of the theorems includes both a frequency domain test and an equivalent state-space test. In addition, robust H_2 performance bounds corresponding to the state-space tests are presented.

Theorem 1 (Small Gain Theorem[7]). If $(1/\gamma)\bar{G}(s)$ is strictly bounded real, then the negative feedback interconnection of $\bar{G}(s)$ and F is asymptotically stable for all $F \in F_\gamma$. Equivalently, if for any symmetric, positive definite $\bar{R}$ there exists a positive scalar α and nonnegative definite P satisfying

$$0 = \bar{A}^T P + P\bar{A} + \gamma^{-2} P\bar{B}_0 \bar{B}_0^T P + \bar{C}_0^T \bar{C}_0 + \alpha\bar{R} \tag{14}$$

then the uncertain system (3) is asymptotically stable for all $F \in F_\gamma$. In this case, for all $F \in F_\gamma$,

$$J(F) \leq \bar{J}(\alpha) = (1/\alpha)\text{tr}(P\bar{V}) \tag{15}$$

Theorem 2 (Positivity Theorem[7]). If $M^{-1} + \bar{G}(s)$ is strongly positive real, then the negative feedback interconnection of $\bar{G}(s)$ and F is asymptotically stable for all $F \in F_M^+$. Equivalently, if for any symmetric, positive definite $\bar{R}$ there exists a positive scalar α and nonnegative definite P satisfying

$$0 = \bar{A}^T P + P\bar{A} + \frac{1}{2}(\bar{C}_0 - \bar{B}_0^T P)^T M^{-1}(\bar{C}_0 - \bar{B}_0^T P) + \alpha\bar{R} \tag{16}$$

then the uncertain system (3) is asymptotically stable for all $F \in F_M^+$. In this case, for all $F \in F_M^+$,

$$J(F) \leq \bar{J}(\alpha) = (1/\alpha)\text{tr}(\bar{P}\bar{V}) \tag{17}$$

Theorem 3 (Popov Theorem[6,7]). If there exists a non-negative-definite diagonal matrix N such that $M^{-1} + (I + Ns)\bar{G}(s)$ is strongly positive real, then the negative feedback interconnection of $\bar{G}(s)$ and F is asymptotically stable for all $F \in F_M^+$. Equivalently, if for any symmetric, positive definite $\bar{R}$ there exists a nonnegative-definite diagonal matrix N, a positive scalar α and nonnegative-definite P satisfying

$$0 = \bar{A}^T P + P\bar{A} + (\bar{C}_0 + N\bar{C}_0\bar{A} - \bar{B}_0^T P)^T[(M^{-1} + N\bar{C}_0\bar{B}_0)$$
$$+ (M^{-1} + N\bar{C}_0\bar{B}_0)^T]^{-1}(\bar{C}_0 + N\bar{C}_0\bar{A} - \bar{B}_0^T P) + \alpha\bar{R} \tag{18}$$

then the uncertain system (3) is asymptotically stable for all $F \in F_M^+$. In this case, for all $F \in F_M^+$,

$$J(F) \leq \bar{J}(\alpha, N) = (1/\alpha)\text{tr}((P + \bar{C}_0^T MN\bar{C}_0)\bar{V}) \tag{19}$$

Remark 1. Theorem 2 may be considered a special case of Theorem 3 with $N = 0$.

Remark 2. In each of the three theorems the requirement that $\bar{R}$ be positive definite can be relaxed. In particular, $\bar{R}$ is

allowed to be nonnegative definite as long as the pair $(\bar{A}, \bar{E})$ is detectable where $\bar{E}$ satisfies $\bar{E}^T\bar{E} = \bar{R}$.

Remark 3. For the case of scalar uncertainty F (i.e., $m_0 = 1$), the frequency domain tests given in the three theorems have easy-to-implement graphical frequency domain interpretations.[2]

Remark 4. As shown in Ref. 7, the Lyapunov function that establishes robust stability of the negative feedback interconnection of $\tilde{G}(s)$ and F in Theorems 1 and 2 is a fixed Lyapunov function of the form $V(\tilde{x}) = \tilde{x}^T P\tilde{x}$ where P satisfies Eqs. (14) and (16), respectively. On the other hand, the Lyapunov function that establishes robust stability of the negative feedback interconnection of $\tilde{G}(s)$ and F in Theorem 3 is a parameter-dependent Lyapunov function; that is, it is a function of the uncertain parameters and has the form $V(\tilde{x}) = \tilde{x}^T P\tilde{x} + \tilde{x}^T\tilde{C}_0^T FN\tilde{C}_0\tilde{x}$ where P satisfies Eq. (18).

Remark 5. Note that the Popov multiplier N can be a negative-definite diagonal matrix that in the single-input/single-output (SISO) case simply corresponds to a Popov line in the Popov plane with a negative slope.[2] In this case, we note that the candidate Lyapunov function has the form $V(\tilde{x}) = \tilde{x}^T P\tilde{x} - \tilde{x}^T\tilde{C}_0^T F\hat{N}\tilde{C}_0\tilde{x}$, where $\hat{N} > 0$. Hence, it is necessary to check a posteriori the positive definiteness of $V(\tilde{x})$ for all $F \in F_M^+$ to insure that $V(\tilde{x})$ is a Lyapunov function.

Remark 6. An alternative statement of Theorem 3 that directly captures uncertainty $F \in F_M^+ \cup F_M^-$ can be obtained by considering the multivariable shifted Popov theorem.[9] Specifically, this case corresponds to replacing M with $2M$ and $\bar{A}$ with $\bar{A} - \bar{B}_0 M\bar{C}_0$ in Theorem 3. In this case the frequency domain interpretation for the case of scalar uncertainty involves a family of frequency-dependent off-axis circles in the Nyquist plane. The circle centers vary as a function of the phase of the Popov multiplier, but each has the same real axis intercepts at $\pm M^{-1}$. For further details see Refs. 9–11.

III. Benchmark Two-Mass/Spring Example

Consider the two-mass/spring system shown in Fig. 2 with uncertain stiffness k. A control force acts on body 1, and the position of body 2 is measured, resulting in a noncollocated control problem. Here, we consider controller 1 of Ref. 8, which was designed for problem 1 of a benchmark problem[12] using the maximum entropy robust control design technique. The controller transfer function given by

$$G_c(s) = \frac{194390(s + 0.33679)[(s - 0.11735)^2 + 0.90996^2]}{(s + 81.438)(s + 131.04)[(s + 2.9049)^2 + 1.8615^2]} \quad (20)$$

was designed so that the closed-loop system is robust with respect to perturbations in the nominal value of the stiffness k (i.e., $k = k_{nom}$). The exact stiffness stability region over which the system will remain stable was computed by a simple search and is given by

$$0.4458 < k < 2.0661 \quad (21)$$

Next, using a graphical approach and the state-space Riccati equation approach, we apply small gain analysis, positivity analysis, and Popov analysis to determine the stiffness stability regions predicted by each of these tests. Each of these tests is related to the previous test and is guaranteed to be less conservative.

We begin by constructing the uncertainty feedback system that will be used in each of the tests. The open-loop plant (for $m_1 = m_2 = 1$) is given by

$$\dot{x}(t) = A(k)x(t) + Bu(t) + D_1 w(t) \quad (22a)$$

$$y(t) = Cx(t) + D_2 w(t) \quad (22b)$$

$$z(t) = E_1 x(t) \quad (22c)$$
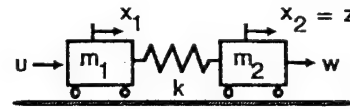
Fig. 2 Benchmark two-mass/spring system for robust control design and analysis.

where

$$A(k) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -k & k & 0 & 0 \\ k & -k & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

$$D_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}, \quad C = E_1 = [0\ 1\ 0\ 0], \quad D_2 = [0\ 1] \quad (23)$$

The H_2 cost functional under consideration is defined with respect to the transfer function between the disturbance $w(t)$ and the performance vector $z(t) + E_2 u(t)$, where $E_2 = \sqrt{10^{-3}}$. The perturbation in $A(k)$ due to a change in the stiffness element k from the nominal value k_{nom} is given by

$$A(k) - A(k_{nom}) \triangleq \Delta A = -B_0 \Delta k C_0 \quad (24)$$

where

$$B_0 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}, \quad C_0 = [1\ -1\ 0\ 0] \quad (25)$$

In the subsequent analysis we will choose $k_{nom} = 0.6$ since the controller (20) was developed under this assumption.

Let the triple (A_c, B_c, C_c) denote the state-space representation of the controller (20). Then, assuming negative feedback, the closed-loop state matrix is given by

$$\bar{A}(k) = \begin{bmatrix} A(k) & -BC_c \\ B_c C & A_c \end{bmatrix}. \quad (26)$$

In addition, $\bar{R}$ and $\bar{V}$ are given by Eqs. (10) and (11) where

$$\bar{E} = [E_1\ -E_2 C_c], \quad \bar{D} = \begin{bmatrix} D_1 \\ B_c D_2 \end{bmatrix} \quad (27)$$

Next, define

$$\bar{B}_0 = \begin{bmatrix} B_0 \\ 0_{4 \times 1} \end{bmatrix}, \quad \bar{C}_0 = [C_0\ \ 0_{1 \times 4}] \quad (28)$$

and recall

$$\bar{G}(s) \sim \left[\begin{array}{c|c} \bar{A} & \bar{B}_0 \\ \hline \bar{C}_0 & 0 \end{array} \right] \quad (29)$$

Then, the plant uncertainty Δk can be represented by the fictitious feedback loop shown in Fig. 1 with $F = \Delta k$. Notice that with this stiffness uncertainty $\bar{C}_0\bar{B}_0 = 0$, which holds for any state-space realization of the system.

IV. Frequency Domain Graphical Analysis of the Benchmark System

In this section we apply the frequency domain tests described in the three theorems of Sec. II to determine $\overline{\Delta k}$ (positive) and $\underline{\Delta k}$ (negative) such that stability is guaranteed for

$$k_{nom} + \underline{\Delta k} < k < k_{nom} + \overline{\Delta k} \quad (30)$$

Since the uncertainty is scalar, we will first use the graphical techniques derived from the frequency domain tests. These graphical tests originally appeared in Refs. 13 and 14 and are included here for comparison with the results based on the state-space formulations.

Small Gain Analysis

Small gain analysis requires considering the Nyquist diagram of $\bar{G}(s)$. The smallest circle centered at the origin that completely encompasses the Nyquist diagram, $\text{Im}[\bar{G}(j\omega)]$ vs $\text{Re}[\bar{G}(j\omega)]$ for all ω, (without touching it) is then drawn. The intersection of this circle with the negative real axis is given by $-1/\Delta k$, and the intersection with the positive real axis is given by $-1/\underline{\Delta k}$. This analysis is shown in Fig. 3. It follows that $\Delta k = 0.1497$ and $\underline{\Delta k} = -0.1497$. Hence, using small gain analysis, stability is guaranteed for

$$0.4503 < k < 0.7497 \tag{31}$$
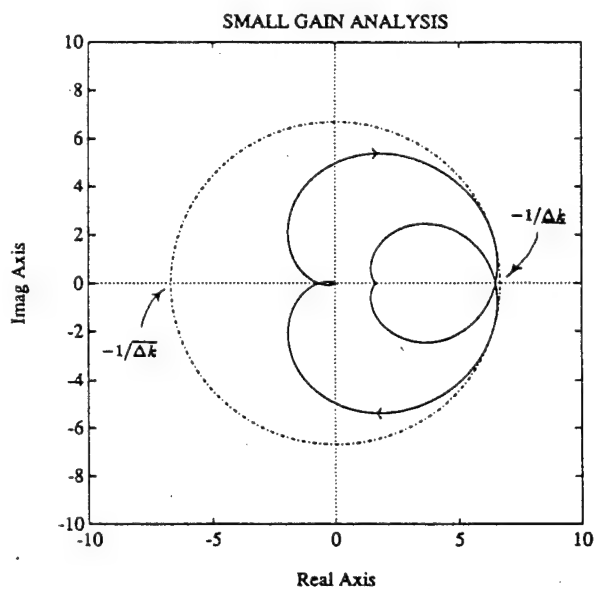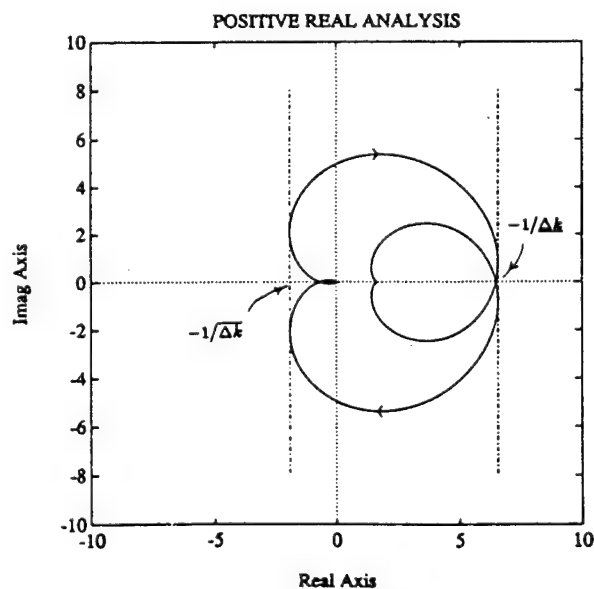
Note that since the Δk uncertainty block is composed of a single scalar, this result is equivalent to the complex structured singular value test.[15]

Positivity Analysis

Positivity analysis determines stability bounds by drawing straight-lines that lie to the left or right of the Nyquist diagram (without touching it). For the Nyquist diagram of $\bar{G}(s)$, the intersection of the line to the left of the Nyquist plot with the negative real axis equals $-1/\Delta k$. The intersection of the line to the right of the Nyquist plot with the positive real axis equals $-1/\underline{\Delta k}$. This analysis is shown in Fig. 4. It follows that $\Delta k = 0.5278$ and $\underline{\Delta k} = -0.1523$. Hence, using positivity analysis, stability is guaranteed for

$$0.4477 < k < 1.1278 \tag{32}$$

Popov Analysis

Popov analysis is a test that determines a stability bound from a modified Nyquist diagram, namely, the Popov plot, $\omega \, \text{Im}[\bar{G}(j\omega)]$ vs $\text{Re}[\bar{G}(j\omega)]$ for $\omega \geq 0$. This analysis requires finding lines (Popov lines) that intersect the negative or positive real axis at a point that is to the left of the Popov plot but as close to the origin as possible. The slopes of these lines are $-1/N$ and $-1/\underline{N}$ where N and $\underline{N}$ are the Popov multipliers. The Popov test is equivalent to the positive real test if the lines are chosen to be vertical. For the Popov diagram of $\bar{G}(s)$, the intersection of the line to the left of the Popov plot with the negative real axis equals $-1/\Delta k$. The intersection of the line to the right of the Popov plot with the positive real axis equals $-1/\underline{\Delta k}$. This analysis is shown in Fig. 5. It follows that $\Delta k = 1.4661$ and $\underline{\Delta k} = 0.1542$, and the corresponding Popov multipliers are, respectively, $\bar{N}^* = 0.7999$ and $\underline{N}^* = -0.2755$. Hence, using Popov analysis, one guarantees stability for

$$0.4458 < k < 2.0661 \tag{33}$$

Note that these bounds are identical to the exact bounds of Eq. (21). Hence, for this example, Popov analysis yields totally nonconservative robust stability results. This is not surprising since, as mentioned in the Introduction, the Popov result is based on a parameter-dependent Lyapunov function that severely restricts the allowable time variation of the uncertain parameters and hence closely approximates real parameter uncertainty within robustness analysis.

Fig. 3   Frequency domain small gain analysis.

Fig. 4   Frequency domain positivity analysis.
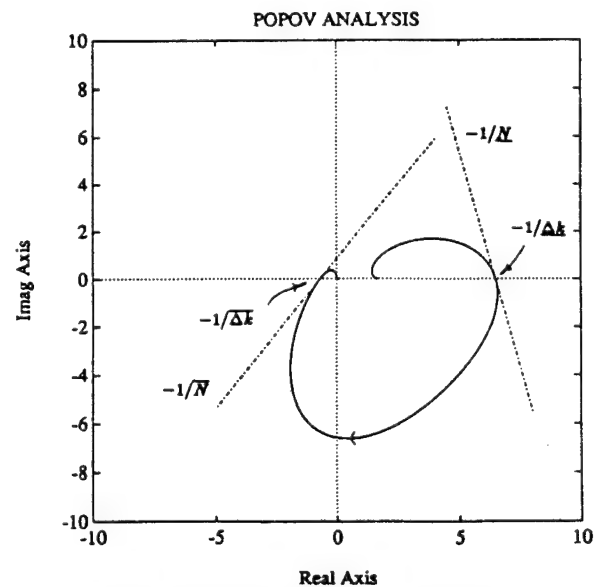
Fig. 5   Frequency domain Popov analysis.

V. State-Space Analysis of the Benchmark System

Continuation (or homotopy) algorithms[16,17] are effective techniques for solving systems of nonlinear algebraic equations and have found increasing engineering applications (see, for example, Refs. 17–19). In this section, we develop continuation algorithms that implement the state-space analysis results described in the three theorems of Sec. II. We restrict ourselves to the case of scalar uncertainty (i.e., F is a scalar) with $\tilde{C}_0\tilde{B}_0 = 0$ (which applies to the benchmark system). In addition, the exposition is focused on implementing state-space Popov analysis since this case is the most complex. The algorithms for small gain and positivity analyses are very similar and hence are only briefly discussed. The results of applying these algorithms to the benchmark problem are subsequently presented.

Each of the algorithms is based on optimizing the cost upper bounds $\bar{J}$ of Eqs. (15), (17), and (19). At this point we focus attention on the upper bound, Eq. (19), of the Popov theorem, rewritten here for all $F \in F_M^+$ as

$$\bar{J}(\alpha, N) \triangleq (1/\alpha)\mathrm{tr}((P + \tilde{C}_0^T MN\tilde{C}_0)\bar{V}) \tag{34}$$

where, for $\tilde{C}_0\tilde{B}_0 = 0$, P is given by

$$0 = \tilde{A}^T P + P\tilde{A} + (\tfrac{1}{2})(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \tilde{B}_0^T P)^T M(\tilde{C}_0 \\ + N\tilde{C}_0\tilde{A} - \tilde{B}_0^T P) + \alpha\tilde{R} \tag{35}$$

The algorithm under consideration will be based on finding scalars α and N that satisfy

$$0 = \frac{\partial \bar{J}}{\partial \alpha} = \mathrm{tr}(Q\tilde{R} - \frac{1}{\alpha^2}(P\bar{V} + \tilde{C}_0^T MN\tilde{C}_0\bar{V})) \tag{36}$$

$$0 = \frac{\partial \bar{J}}{\partial N} = \frac{1}{\alpha}M\tilde{C}_0\bar{V}\tilde{C}_0^T + M(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \tilde{B}_0^T P)Q\tilde{A}^T\tilde{C}_0^T \tag{37}$$

where Q satisfies

$$0 = (A - \tfrac{1}{2}\tilde{B}_0 M(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \tilde{B}_0^T P))Q \\ + Q(A - \tfrac{1}{2}\tilde{B}_0 M(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \tilde{B}_0^T P))^T + (1/\alpha)\bar{V} \tag{38}$$

Continuation Map for Popov Analysis

To define the continuation map we assume that the uncertainty parameter M is a function of the continuation parameter $\lambda \in [0, 1]$. In particular, it is assumed that

$$M(\lambda) = M_0 + \lambda(M_f - M_0) \tag{39}$$

Note that, at $\lambda = 0$, $M(\lambda) = M_0$, whereas at $\lambda = 1$, $M(\lambda) = M_f$. The continuation map is defined as the gradient of the upper bound on the cost for the uncertainty parameter $M(\lambda)$. In particular,

$$H(\theta, \lambda) \triangleq \begin{bmatrix} H_1(\theta, \lambda) \\ H_2(\theta, \lambda) \end{bmatrix} \tag{40}$$

where

$$\theta \triangleq \begin{bmatrix} \alpha \\ N \end{bmatrix} \tag{41}$$

$$H_1(\theta, \lambda) \triangleq \mathrm{tr}(Q(\theta, \lambda)\tilde{R} - \frac{1}{\alpha(\lambda)^2}(P(\theta, \lambda)\bar{V} + \tilde{C}_0^T M(\lambda)N(\lambda)\tilde{C}_0\bar{V})) \tag{42}$$

$$H_2(\theta, \lambda) \triangleq \frac{1}{\alpha(\lambda)}M(\lambda)\tilde{C}_0\bar{V}\tilde{C}_0^T \\ + M(\lambda)(\tilde{C}_0 + N(\lambda)\tilde{C}_0\tilde{A} - \tilde{B}_0^T P(\theta, \lambda))Q(\theta, \lambda)\tilde{A}^T\tilde{C}_0^T \tag{43}$$

and

$$0 = \tilde{A}^T P(\theta, \lambda) + P(\theta, \lambda)\tilde{A} \\ + \tfrac{1}{2}(\tilde{C}_0 + N(\lambda)\tilde{C}_0\tilde{A} - \tilde{B}_0^T P(\theta, \lambda))^T M(\lambda) \\ \cdot (\tilde{C}_0 + N(\lambda)\tilde{C}_0\tilde{A} - \tilde{B}_0^T P(\theta, \lambda)) + \alpha(\lambda)\tilde{R} \tag{44}$$

$$0 = [\tilde{A} - \tfrac{1}{2}\tilde{B}_0 M(\lambda)(\tilde{C}_0 + N(\lambda)\tilde{C}_0\tilde{A} - \tilde{B}_0^T P(\theta, \lambda))]Q(\theta, \lambda) \\ + Q(\theta, \lambda)[\tilde{A} - \tfrac{1}{2}\tilde{B}_0 M(\lambda)(\tilde{C}_0 + N(\lambda)\tilde{C}_0\tilde{A} \\ - \tilde{B}_0^T P(\theta, \lambda))]^T + \frac{1}{\alpha(\lambda)}\bar{V} \tag{45}$$

The continuation curve is defined by

$$0 = H(\theta, \lambda), \qquad \lambda \in [0, 1] \tag{46}$$

Jacobian of the Continuation Map for Popov Analysis

The algorithm requires computation of $\nabla H(\theta, \lambda)^T$, the Jacobian of $H(\theta, \lambda)$. Note that

$$\nabla H(\theta, \lambda)^T = [H_\theta \ H_\lambda] \tag{47}$$

where

$$H_\theta \triangleq \frac{\partial H}{\partial \theta} \tag{48a}$$

$$H_\lambda \triangleq \frac{\partial H}{\partial \lambda} \tag{48b}$$

Expressions for H_θ and H_λ are given in the Appendix.

Outline of the Continuation Algorithm for Popov Analysis

Step 1. Initialize $loop = 0$, $\lambda = 0$, $\Delta\lambda \in [0, 1]$, $\theta^T = [1\ 0]$ (i.e., $\alpha = 1$, $N = 0$).
Step 2. Let $loop = loop + 1$. If $loop = 1$, then go to step 4. Otherwise, continue.
Step 3. Advance the homotopy parameter and predict the corresponding parameter vector θ as follows.
3a. Let $\lambda_0 = \lambda$.
3b. Let $\lambda = \lambda_0 + \Delta\lambda$.
3c. Compute $H_\theta(\theta, \lambda)$ and $H_\lambda(\theta, \lambda)$. Then compute $\theta'_p(\lambda_0)$ using

$$\theta'_p(\lambda_0) = -[H_\theta(\theta, \lambda)]^{-1}H_\lambda(\theta, \lambda_0) \tag{49}$$

3d. Predict $\theta(\lambda)$ using $\theta(\lambda) = \theta(\lambda_0) + \Delta\lambda\theta'_p(\lambda_0)$.
3e. If $\|H(\theta, \lambda)\|$ satisfies some preassigned prediction tolerance, then continue. Otherwise, reduce $\Delta\lambda$ and go to step 3b.
Step 4. Correct the current approximation $\theta(\lambda)$ as follows.
4a. Compute $H(\theta, \lambda)$ and $H_\theta(\theta, \lambda)$.
4b. Correct $\theta(\lambda)$ using $\theta(\lambda) \leftarrow \theta(\lambda) - [H_\theta(\theta, \lambda)]^{-1}H(\theta, \lambda)$.
4c. If $\|H(\theta, \lambda)\|$ satisfies some preassigned tolerance, then continue. Otherwise, go to step 4a.
4d. If $P(\lambda)$ is not nonnegative definite, then go to step 5, since stability is only guaranteed for $M = M(\lambda_0)$. Otherwise, continue.
4e. Compute the upper bound $\bar{J}(\theta)$.
4f. If $\lambda = 1$, then continue. Otherwise, go to step 2.
Step 5. Stop.

Continuation Algorithm for Positivity Analysis

Recall that positivity analysis is a special case of Popov analysis (with $N \triangleq 0$). Hence, positivity analysis is implemented using the algorithm for Popov analysis with N constrained to zero.

Continuation Algorithm for Small Gain Analysis

For small gain analysis we consider the upper bound $\bar{J}(\alpha)$ of Eq. (15), rewritten here as

$$\bar{J}(\alpha) \triangleq (1/\alpha)\mathrm{tr}(P\bar{V}) \tag{50}$$

where

$$0 = \bar{A}^T P + P\bar{A} + M^{-2}P\bar{B}_0\bar{B}_0^T P + \bar{C}_0^T\bar{C}_0 + \alpha\bar{R} \tag{51}$$

The algorithm is based on finding a scalar α such that

$$0 = \frac{\partial \bar{J}}{\partial \alpha} = \mathrm{tr}\left(Q\bar{R} - \frac{1}{\alpha^2}P\bar{V}\right) \tag{52}$$

where Q satisfies

$$0 = (\bar{A} + M^{-2}\bar{B}_0\bar{B}_0^T P)Q + Q(\bar{A} + M^{-2}\bar{B}_0\bar{B}_0^T P)^T + (1/\alpha)\bar{V} \tag{53}$$

It is assumed that $M(\lambda)$ is as given by Eq. (39) and the continuation map is defined as

$$H(\theta, \lambda) = \mathrm{tr}\left(Q(\theta, \lambda)\bar{R} - \frac{1}{\alpha^2}P(\theta, \lambda)\bar{V}\right) \tag{54}$$

where

$$\theta \triangleq \alpha \tag{55}$$
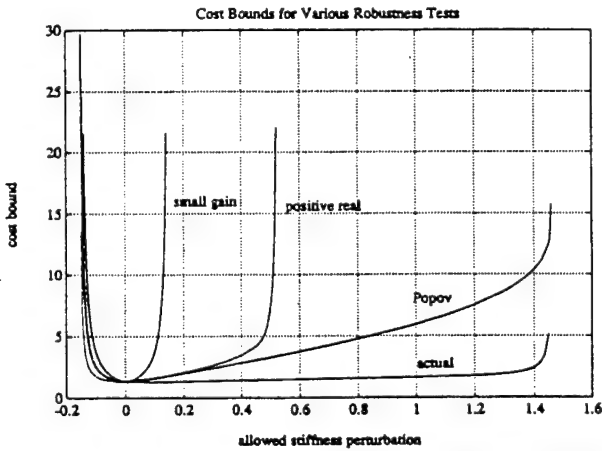
Fig. 6   Performance bounds for the small gain, positivity, and Popov tests.
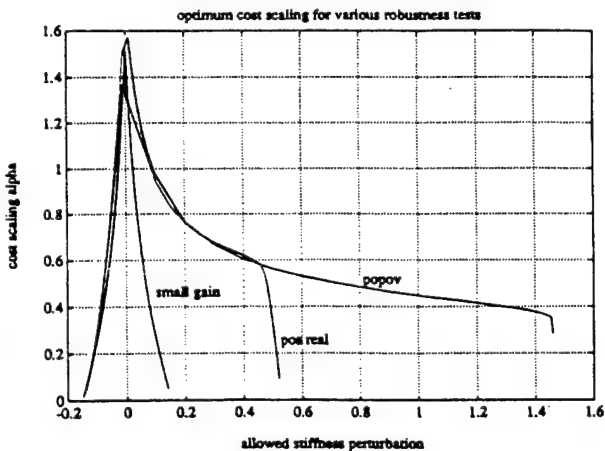
Fig. 7   Optimal α for the small gain, positivity, and Popov tests.
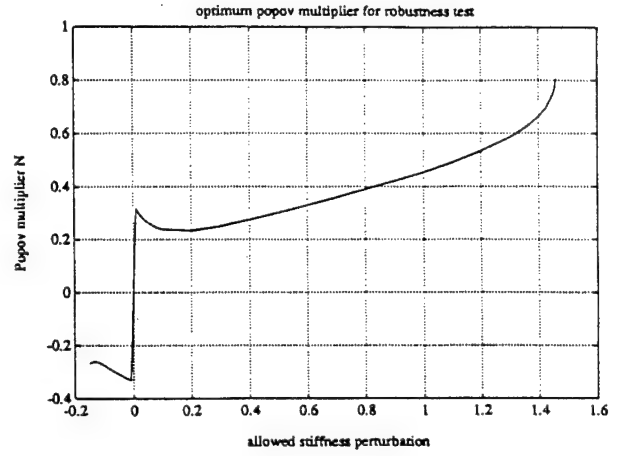
Fig. 8   Optimal N for the Popov test.

$$0 = \bar{A}^T P(\theta, \lambda) + P(\theta, \lambda)\bar{A}$$
$$+ M(\lambda)^{-2}P(\theta, \lambda)\bar{B}_0\bar{B}_0^T P(\theta, \lambda) + \bar{C}_0^T\bar{C}_0 + \alpha\bar{R} \tag{56}$$

$$0 = (\bar{A} + M(\lambda)^{-2}\bar{B}_0\bar{B}_0^T P(\theta, \lambda))Q(\theta, \lambda)$$
$$+ Q(\theta, \lambda)(\bar{A} + M(\lambda)^{-2}\bar{B}_0\bar{B}_0^R P(\theta, \lambda))^T + (1/\alpha)\bar{V} \tag{57}$$

The continuation curve is defined by

$$0 = H(\theta, \lambda), \qquad \lambda \in [0, 1] \tag{58}$$

Expressions for the Hessian H_θ and H_λ are given in the Appendix.

The outline of the continuation algorithm for small gain analysis is identical to that given for Popov analysis. Because of this, no further discussion is needed.

Analysis of the Benchmark Problem

When the continuation algorithms for small gain, positivity, and Popov analysis are applied to the benchmark problem, the performance curves shown in Fig. 6 result. As expected, Popov analysis yields less conservative results than the positivity and small gain tests. The robust stability bounds $\overline{\Delta k}$ (positive) and $\underline{\Delta k}$ (negative) obtained from the state-space tests are identical to those obtained from the frequency domain tests of Sec. IV. The optimal α for each test is shown in Fig. 7 as a function of M. The optimal N for the Popov test is shown in Fig. 8. Note that as M approaches its supremum and infimum, N converges, respectively, to $\overline{N}^*$ and $\underline{N}^*$ obtained from the graphical test.

VI.   Conclusions

This paper has discussed the small gain, positivity, and Popov tests and applied both the (graphical) frequency domain version of each test and the corresponding state-space test to a benchmark problem. The frequency domain tests and the state-space tests were seen to give identical results for robust stability, and the Popov test was completely nonconservative in its robustness predictions. The state-space tests also yielded robust H_2 performance bounds and were implemented using continuation algorithms. The algorithms developed here only apply to the special case of scalar uncertainty and the algorithm for the Popov test further requires that a certain product (related to the uncertainty pattern) is zero. Future work will involve the development of more general numerical algorithms.

Appendix: Jacobian Expressions for the Popov and Small Gain Tests

In this Appendix we show how to compute the Jacobian of the homotopy map $H(\theta, \lambda)$ for both the Popov and small gain

tests. We first recall that the Jacobian $\nabla H(\theta, \lambda)$ is defined by

$$\nabla H(\theta, \lambda)^T \triangleq [H_\theta, H_\lambda] \tag{A1}$$

where

$$H_\theta \triangleq \frac{\partial H}{\partial \theta}, \qquad H_\lambda \triangleq \frac{\partial H}{\partial \lambda} \tag{A2}$$

Since $H(\theta, \lambda)$ corresponds to the upper bounds of the cost corresponding to $M(\lambda)$, H_θ is the corresponding Hessian.

Jacobian Expressions for Popov Analysis

$$H_\theta = \begin{bmatrix} \dfrac{\partial}{\partial \alpha} H_1(\theta, \lambda) & \dfrac{\partial}{\partial N} H_1(\theta, \lambda) \\[3mm] \text{SYM} & \dfrac{\partial}{\partial N} H_2(\theta, \lambda) \end{bmatrix} \tag{A3}$$

where

$$\frac{\partial}{\partial \alpha} H_1 = \text{tr}\left(\frac{\partial Q}{\partial \alpha} \bar{R} - \frac{1}{\alpha^2} \frac{\partial P}{\partial \alpha} \bar{V} + \frac{2}{\alpha^3} (P\bar{V} + \tilde{C}_0^T MN\tilde{C}_0 \bar{V}) \right) \tag{A4}$$

$$\frac{\partial}{\partial N} H_2 = \left(-M\bar{B}_0^T \frac{\partial P}{\partial N} Q + M(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \bar{B}_0^T P)\frac{\partial Q}{\partial N} \right.$$
$$\left. + M\tilde{C}_0\tilde{A}Q)\tilde{A}^T\tilde{C}_0^T \right) \tag{A5}$$

$$\frac{\partial}{\partial N} H_1 = \text{tr}\left(\frac{\partial Q}{\partial N} \bar{R} - \frac{1}{\alpha^2} \frac{\partial P}{\partial N} \bar{V} - \frac{1}{\alpha^2} \tilde{C}_0^T M\tilde{C}_0 \bar{V} \right) \tag{A6}$$

and $\partial P/\partial \alpha$, $\partial P/\partial N$, $\partial Q/\partial \alpha$, and $\partial Q/\partial N$ satisfy

$$0 = \tilde{A}_a^T \frac{\partial P}{\partial \alpha} + \frac{\partial P}{\partial \alpha} \tilde{A}_a + \bar{R} \tag{A7}$$

$$0 = \tilde{A}_a^T \frac{\partial P}{\partial N} + \frac{\partial P}{\partial N} \tilde{A}_a + \tfrac{1}{2}[(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \bar{B}_0^T P)^T M\tilde{C}_0\tilde{A}$$
$$+ \tilde{A}^T\tilde{C}_0^T M(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - B_0^T P)] \tag{A8}$$

$$0 = \tilde{A}_a \frac{\partial Q}{\partial \alpha} + \frac{\partial Q}{\partial \alpha} \tilde{A}_a^T + \left(\frac{1}{2}\bar{B}_0 M\bar{B}_0^T \frac{\partial P}{\partial \alpha} \right) Q$$
$$+ Q\left(\frac{1}{2}\bar{B}_0 M\bar{B}_0^T \frac{\partial P}{\partial \alpha} \right)^T - \frac{1}{\alpha^2}\bar{V} \tag{A9}$$

$$0 = \tilde{A}_a \frac{\partial Q}{\partial N} + \frac{\partial Q}{\partial N} \tilde{A}_a^T + \left(\frac{1}{2}\bar{B}_0 M\bar{B}_0^T \frac{\partial P}{\partial N} \right) Q + Q\left(\frac{1}{2}\bar{B}_0 M\bar{B}_0^T \frac{\partial P}{\partial N} \right)^T$$
$$- \left(\frac{1}{2}\bar{B}_0 M\tilde{C}_0\tilde{A} \right) Q - Q\left(\frac{1}{2}\bar{B}_0 M\tilde{C}_0\tilde{A} \right)^T \tag{A10}$$

where

$$\tilde{A}_a \triangleq \tilde{A} - \tfrac{1}{2}\bar{B}_0 M(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \bar{B}_0^T P) \tag{A11}$$

Similarly, H_λ is given by

$$H_\lambda = \begin{bmatrix} \dfrac{\partial}{\partial \lambda} H_1(\theta, \lambda) \\[3mm] \dfrac{\partial}{\partial \lambda} H_2(\theta, \lambda) \end{bmatrix} \tag{A12}$$

where

$$\frac{\partial}{\partial \lambda} H_1 = \text{tr}\left(\frac{\partial Q}{\partial \lambda} \bar{R} - \frac{1}{\alpha^2}\left(\frac{\partial P}{\partial \lambda} \bar{V} + \tilde{C}_0^T (M_f - M_0)N\tilde{C}_0\bar{V} \right) \right) \tag{A13}$$

$$\frac{\partial}{\partial \lambda} H_2 = \frac{1}{\alpha}(M_f - M_0)\tilde{C}_0\bar{V}\tilde{C}_0^T$$
$$+ \left[(M_f - M_0)(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \bar{B}_0^T P)Q - M\bar{B}_0^T\frac{\partial P}{\partial \lambda}Q \right.$$
$$\left. + M(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \bar{B}_0^T P)\frac{\partial Q}{\partial \lambda} \right] \tilde{A}^T\tilde{C}_0^T \tag{A14}$$

and $\partial P/\partial \lambda$ and $\partial Q/\partial \lambda$ satisfy

$$0 = \tilde{A}_a^T \frac{\partial P}{\partial \lambda} + \frac{\partial P}{\partial \lambda} \tilde{A}_a$$
$$+ \tfrac{1}{2}(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \bar{B}_0^T P)(M_f - M_0)(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \bar{B}_0^T P) \tag{A15}$$

$$0 = \tilde{A}_a \frac{\partial Q}{\partial \lambda} + \frac{\partial Q}{\partial \lambda} \tilde{A}_a^T$$
$$+ \left(\frac{1}{2}\bar{B}_0 M\bar{B}_0^T\frac{\partial P}{\partial \lambda} \right)Q + Q\left(\frac{1}{2}\bar{B}_0 M\bar{B}_0^T\frac{\partial P}{\partial \lambda} \right)^T$$
$$- \left[\frac{1}{2}\bar{B}_0(M_f - M_0)(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \bar{B}_0^T P) \right]Q$$
$$- Q\left[\frac{1}{2}\bar{B}_0(M_f - M_0)(\tilde{C}_0 + N\tilde{C}_0\tilde{A} - \bar{B}_0^T P) \right] \tag{A16}$$

Jacobian Expressions for Small Gain Analysis

The Hessian $H_\theta(\triangleq \partial H/\partial \theta)$ is given by

$$H_\theta = \text{tr}\left(\frac{\partial Q}{\partial \alpha} \bar{R} - \frac{1}{\alpha^2} \frac{\partial P}{\partial \alpha} \bar{V} + \frac{2}{\alpha^3}P\bar{V} \right) \tag{A17}$$

where $\partial P/\partial \alpha$ and $\partial Q/\partial \alpha$ satisfy

$$0 = \tilde{A}^T\frac{\partial P}{\partial \alpha} + \frac{\partial P}{\partial \alpha}\tilde{A} + \bar{R} \tag{A18}$$

$$0 = (\tilde{A} + M^{-2}\bar{B}_0\bar{B}_0^T P)\frac{\partial Q}{\partial \alpha} + \frac{\partial Q}{\partial \alpha}(\tilde{A} + M^{-2}\bar{B}_0\bar{B}_0^T P)^T$$
$$+ M^{-2}\bar{B}_0\bar{B}_0^T\frac{\partial P}{\partial \alpha}Q + \left(M^{-2}\bar{B}_0\bar{B}_0^T\frac{\partial P}{\partial \alpha}Q \right)^T - \frac{1}{\alpha^2}\bar{V} \tag{A19}$$

Similarly, $H_\lambda(\triangleq \partial H/\partial \lambda)$ is given by

$$H_\lambda = \text{tr}\left(\frac{\partial Q}{\partial \lambda}\bar{R} - \frac{1}{\alpha^2}\frac{\partial P}{\partial \lambda}\bar{V} \right) \tag{A20}$$

where $\partial P/\partial \lambda$ and $\partial Q/\partial \lambda$ satisfy

$$0 = (\tilde{A} + M^{-2}\bar{B}_0\bar{B}_0^T P)^T\frac{\partial P}{\partial \lambda} + \frac{\partial P}{\partial \lambda}(\tilde{A} + M^{-2}\bar{B}_0\bar{B}_0^T P)$$
$$- 2M^{-3}(M_f - M_0)P\bar{B}_0\bar{B}_0^T P \tag{A21}$$

$$0 = (\tilde{A} + M^{-2}\bar{B}_0\bar{B}_0^T P)\frac{\partial Q}{\partial \lambda} + \frac{\partial Q}{\partial \lambda}(\tilde{A} + M^{-2}\bar{B}_0\bar{B}_0^T P)^T$$
$$+ M^{-2}\bar{B}_0\bar{B}_0^T\frac{\partial P}{\partial \lambda}Q + \left(M^{-2}\bar{B}_0\bar{B}_0^T\frac{\partial P}{\partial \lambda}Q \right)^T$$
$$- 2M^{-3}(M_f - M_0)(\bar{B}_0\bar{B}_0^T PQ + QP\bar{B}_0\bar{B}_0^T) \tag{A22}$$

References

[1]Zames, G., "On the Input-Output Stability of Time-Varying Nonlinear Feedback Systems, Part I: Conditions Derived Using Concepts of Loop Gain, Conicity, and Positivity," *IEEE Transactions on Automatic Control*, Vol. AC-11, April 1966, pp. 228–238.

[2]Narendra, K. S., and Taylor, J. H., *Frequency Domain Criteria for Absolute Stability*, Academic Press, New York, 1973.

[3]Francis, B. A., and Doyle, J. C., "Linear Control Theory with an H_∞ Optimality Criterion," *SIAM Journal of Control and Optimization*, Vol. 25, July 1987, pp. 815–844.

[4]Francis, B. A., *A Course in H_∞ Control Theory*, Springer-Verlag, New York, 1987.

[5]Khargonekar, P. P., Petersen, I. R., and Zhou, K., "Robust Stabilization of Uncertain Linear Systems: Quadratic Stability and H_∞ Theory," *IEEE Transactions on Automatic Control*, Vol. 35, March 1990, pp. 356–361.

[6]Haddad, W. M., and Bernstein, D. S., "Parameter-Dependent Lyapunov Functions, Constant Real Parameter Uncertainty, and the Popov Criterion in Robust Analysis and Synthesis: Part 1, Part 2," *Proceedings of the IEEE Conference on Decision and Control* (Brighton, England, UK), IEEE, Piscataway, NJ, 1991, pp. 2274–2279 and 2618–2623; also *IEEE Transactions on Automatic Control* (submitted for publication).

[7]Haddad, W. M., and Bernstein, D. S., "Explicit Construction of Quadratic Lyapunov Functions for the Small Gain, Positivity, Circle, and Popov Theorems and Their Application to Robust Stability, Part 1: Continuous-Time Theory, Part 2: Discrete-Time Theory," *International Journal on Robust and Nonlinear Control* (to be published).

[8]Collins, E. G., Jr., King, J. A., and Bernstein, D. S., "Application of Maximum Entropy/Optimal Projection Design Synthesis to a Benchmark Problem," *Journal of Guidance, Control, and Dynamics*, Vol. 15, No. 5, 1992, pp. 1094–1102.

[9]Haddad, W. M., and Bernstein, D. S., "The Multivariable Parabola Criterion for Robust Controller Synthesis: A Riccati Equation Approach," *Journal of Mathematical Systems, Estimation, and Control* (to be published).

[10]How, J. P., and Hall, S. R., "Connections Between the Popov Stability Criterion and Bounds for Real Parametric Uncertainty," *IEEE Transactions on Automatic Control* (submitted for publication).

[11]Haddad, W. M., How, J. P., Hall, S. R., and Bernstein, D. S., "Extensions of Mixed-μ Bounds to Monotonic and Odd Monotonic Nonlinearities Using Absolute Stability Theory," *International Journal of Control* (to be published).

[12]Wie, B., and Bernstein, D. S., "Benchmark Problems for Robust Control Design," *Journal of Guidance, Control, and Dynamics*, Vol. 25, No. 5, 1992, pp. 1057–1059.

[13]Collins, E. G., Jr., Haddad, W. M., and Bernstein, D. S., "Small Gain, Circle, Positivity, and Popov Analysis of a Maximum Entropy Controller for a Benchmark Problem," *Proceedings of the American Control Conference* (Chicago, IL), IEEE, Piscataway, NJ, 1991, pp. 2425–2426.

[14]Haddad, W. M., Collins, E. G., Jr., and Bernstein, D. S., "Robust Stability Analysis Using the Small Gain, Circle, Positivity, and Popov Theorems: A Comparative Study," *IEEE Transactions on Automatic Control Systems Technology*, Dec. 1993.

[15]Doyle, J. C., "Analysis of Feedback Systems with Structured Uncertainties," *IEE Proceedings, Part D*, Vol. 129, Nov. 1982, pp. 242–250.

[16]Watson, L. T., "Globally Convergent Homotopy Algorithms for Nonlinear Systems of Equations," *Nonlinear Dynamics*, Vol. 1, 1990, pp. 143–191.

[17]Richter, S. L., and DeCarlo, R. A., "Continuation Methods: Theory and Applications," *IEEE Transactions on Automatic Control*, Vol. CAS-30, No. 6, 1983, pp. 347–352.

[18]Lefebvre, S., Richter, S., and DeCarlo, R., "A Continuation Algorithm for Eigenvalue Assignment by Decentralized Constant-Output Feedback," *International Journal of Control*, Vol. 41, No. 5, 1985, pp. 1273–1292.

[19]Collins, E. G., Jr., Davis, L. D., and Richter, S., "A Homotopy Algorithm for Maximum Entropy Design," *Journal of Guidance, Control, and Dynamics* (to be published); also *Proceedings of the American Control Conference* (San Francisco, CA), IEEE, Piscataway, NJ, 1993, pp. 1010–1014.

Appendix F:

Frequency Domain Performance Bounds for Uncertain
Positive Real Plants Controlled by Strictly Positive Real Compensators

Frequency Domain Performance Bounds for Uncertain Positive Real Plants Controlled by Strictly Positive Real Compensators

by

David C. Hyland
Harris Corporation
Government Aerospace Systems
Division
MS 19/4849
Melbourne, Florida 32902
(407) 729-2138
FAX: (407) 727-4016

Emmanuel G. Collins, Jr.
Harris Corporation
Government Aerospace Systems
Division
MS 19/4849
Melbourne, Florida 32902
(407) 727-6358
FAX: (407) 727-4016
ecollins@ x102a.ess.harris.com

Wassim M. Haddad
Department of Mechanical and
Aerospace Engineering
Florida Institute of Technology
Melbourne, Florida 32901
(407) 768-8000 Ext.7241
FAX: (407) 984-8461
haddad@ zach.fit.edu

Vijaya S. Chellaboina
Department of Mechanical and
Aerospace Engineering
Florida Institute of Technology
Melbourne, Florida 32901
(407) 768-8000 Ext.7630
FAX: (407) 984-8461
vijaya@ ee.fit.edu

Abstract

An important part of feedback control involves analyzing uncertain systems for robust stability and performance. Many robustness theories consider only stability issues and ignore performance. Most of the performance robustness results that do exist will not always yield finite performance bounds for the case of closed-loop systems consisting of uncertain positive real plants controlled by strictly positive real compensators. These results are obviously conservative since this class of systems is unconditionally stable. This paper uses majorant analysis to develop tests that yield finite performance bounds for the above case. The results are specialized to the case of static, decentralized colocated rate feedback and dynamic colocated rate feedback.

1. Introduction

A central issue in feedback control is the analysis of uncertain systems for robust stability and performance. Hence, considerable effort has been devoted by researchers in control to the development of effective robustness analysis tools. Many of the developments in robustness analysis have focused exclusively on the determination of stability. However, in practical engineering, performance issues are paramount, so that it is important to additionally determine the type of performance degradation that occurs due to the uncertainty in the system modeling. References [1-13] are examples of robustness analysis techniques that do consider performance. A common feature of a class of these results [5-8] is that they rely on majorant bounding techniques [14-16].

Majorant theory was originally developed by Dahlquist to produce bounds for the solutions of systems of differential equations [16]. The corresponding bounding techniques focus on providing upper bounds on subblocks of matrices and inverse matrices. Similar bounding procedures have been used in the work of researchers in large scale systems analysis [17,18]. The more recent results of [5-8] apply majorant techniques to produce robust performance bounds for uncertain linear systems.

In [5-8] performance is measured in basically three ways. References [5] and [6] measure performance in terms of second order statistics. In particular, bounds are obtained on the steady state variances of selected system variables. In [7], performance is expressed in terms of the frequency response of selected system outputs. This result led to a new upper bound for the structured singular value. Finally, [8] considers the transient response of certain system outputs, a performance measure which had not previously been treated in the robustness literature. A common feature of these results and most other robustness results, with the possible exception of methods based on extensions of Popov analysis and parameter-dependent Lyapunov functions [11-13], is that they do *not* predict unconditional stability for feedback systems consisting of a positive real plant controlled by a strictly positive real controller.

This paper uses the logarithmic norm in context of majorant analysis to develop tests

1

for robust stability and performance that predict unconditional stability for the above case and also yield robust performance bounds. As in [1,2,7,10] this paper considers the frequency domain behavior of a given system. The results are specialized to the case of static, decentralized colocated rate feedback and dynamic, colocated rate feedback. The bounds developed here are illustrated with examples chosen from this class of problems and compared with the performance bound obtained in [7] and the performance bound resulting from complex structured singular value analysis [1,2]. It is seen that the new bounds are much less conservative than the alternative bounds.

The paper is organized as follows. Section 2 presents notation and the necessary mathematical foundation. Section 3 gives results relating to strictly positive real feedback of a positive real system. Section 4 develops robust performance bounds for the aforementioned systems. Section 5 specializes the performance bounds to the case of static, decentralized colocated rate feedback. In Section 6 we extend the results of Section 5 to dynamic, centralized output feedback and present a systematic approach for designing strictly positive real compensators. In order to draw comparisions to the robust performance bounds developed in Section 5 and 6, Section 7 presents a brief summary of the results developed in [7] and [1,2] involving an alternative majorant bound and the complex structured singular value bound respectively. Section 8 presents several illustrative examples that demonstrate the effectiveness of the proposed approach. Finally, Section 9 presents conclusions.

2. Notation and Mathematical Preliminaries

In the following notation, the matrices and vectors are in general assumed to be complex.

$\mathbb{R}$	set of real numbers
$\mathbb{C}$	set of complex numbers
I_p	$p \times p$ identity matrix
Z^*	complex conjugate of matrix Z
Z^H	complex conjugate transpose of matrix Z $(= (Z^*)^{\mathrm{T}})$
z_{ij} or Z_{ij}	(i,j) element of matrix Z
$\mathrm{diag}\{z_1, \ldots, z_n\}$	diagonal matrix with listed diagonal elements
$Y \leq\leq Z$	$y_{ij} \leq z_{ij}$ for each i and j, where Y and Z are real matrices with identical dimensions

$\lvert\alpha\rvert$	absolute value of complex scalar α
$\det(Z)$	determinant of square matrix Z
$\lVert x\rVert_2$	Euclidean norm of vector x $(=\sqrt{x^H x})$
$\sigma_{\min}(Z), \sigma_{\max}(Z)$	minimum, maximum singular values of matrix Z
$\lVert Z\rVert_s$	spectral norm of matrix Z $(=\sigma_{\max}(Z))$, subordinate to the Euclidean norm
$\lVert Z\rVert_F$	Frobenius norm of matrix Z $\left(=\sqrt{\sum_i \sum_j z_{ij} z_{ij}^*}\right)$
$\rho(Z)$	spectral radius of a square matrix Z
$\mathcal{L}[Z(t)]$	Laplace transform of $Z(t)$
$\lambda_{\min}(Q), \lambda_{\max}(Q)$	minimum, maximum eigenvalues of the Hermitian matrix Q
$\max\{Y_1,\ldots,Y_n\}$	$=\overline{Y}$ where $\overline{y}_{ij} = \max\{y_{1,ij}, y_{2,ij}, \ldots, y_{n,ij}\}$

Let $A \in \mathbb{C}^{m\times n}$. Then, the *modulus matrix* of A is the $m \times n$ nonnegative matrix

$$|A|_M \triangleq [|a_{ij}|]. \tag{2.1}$$

The modulus matrix is a special case of a *block norm matrix* [14,15].

Let $B \in \mathbb{C}^{n\times p}$. Subsequent analysis will use the following relation

$$|AB|_M \leq\leq |A|_M |B|_M. \tag{2.2}$$

A *majorant* [16] is an element-by-element upper bound for a modulus matrix (or more generally, a block norm matrix). Specifically, $\hat{A}$ is an $m \times n$ *majorant* respectively of $A \in \mathbb{C}^{m\times n}$ if

$$|A|_M \leq\leq \hat{A}. \tag{2.3}$$

Let $Z \in \mathbb{C}^{n\times n}$. Then $\check{Z} \in \mathbb{R}^{n\times n}$ is an $n \times n$ *minorant* [16] of Z if

$$\check{z}_{ii} \leq |z_{ii}|, \tag{2.4a}$$

$$\check{z}_{ij} \leq -|z_{ij}|, \quad i \neq j. \tag{2.4b}$$

Lemma 2.1. Let Z_d and Z_{od} denote respectively the diagonal and off-diagonal components of $Z \in \mathbb{C}^{n\times n}$, such that

$$Z_d = \mathrm{diag}\{z_{ii}\}_{i=1}^n, \quad Z_{od} = Z - Z_d. \tag{2.5}$$

3

Then, if $\check{Z}_{\mathrm{d}}$ is an $n \times n$ minorant of Z_{d} and $\hat{Z}_{\mathrm{od}}$ is a majorant of Z_{od}, $\check{Z}_{\mathrm{d}} - \hat{Z}_{\mathrm{od}}$ is a minorant of Z.

The *logarithmic norm* [16,19] of $Z \in \mathbb{C}^{n \times n}$ with respect to the spectral norm is defined by

$$\tau(Z) \triangleq \lim_{h \to 0+} \frac{\|I + hZ\|_{\mathrm{s}} - 1}{h}, \tag{2.6}$$

or, equivalently [19],

$$\tau(Z) = \frac{1}{2}\lambda_{\max}(Z + Z^H). \tag{2.7}$$

A matrix $P \in \mathbb{R}^{n \times n}$ is an *M-matrix* [20-22] if it has nonpositive off-diagonal elements (i.e., $p_{ij} \leq 0$ for $i \neq j$) and positive principal minors. It has been shown [20-22] that the inverse of an M-matrix is a nonnegative matrix.

The next five lemmas, especially Lemmas 2.4 and 2.6, are key to the development of the robust performance bounds of the following sections. The proofs of these lemmas are based on the relationship between minorants, logarithmic norms, and M-matrices.

Lemma 2.2. Let $Z \in \mathbb{C}^{n \times n}$. Then $\check{Z} \in \mathbb{R}^{n \times n}$ is a $n \times n$ minorant of Z if

$$\check{z}_{ii} \leq \frac{1}{2}(z_{ii} + z_{ii}^{*}), \tag{2.8}$$

or

$$\check{z}_{ii} \leq \frac{1}{2}|\jmath(z_{ii} - z_{ii}^{*})|, \tag{2.9}$$

and

$$\check{z}_{ij} \leq -|z_{ij}|, \quad i \neq j. \tag{2.10}$$

Proof. It follows from equation (3.1) of [16] that $\check{Z}$ is an $n \times n$ minorant of Z if

$$\check{z}_{ii} \leq -\tau(-z_{ii}), \tag{2.11}$$

and (2.10) is satisfied. Substituting (2.7) into (2.11) yields (2.8). Hence $\check{Z}$ satisfying (2.8) and (2.10) is a $n \times n$ minorant of Z.

4

Next recognizing that a minorant of Z^H is also a minorant of $\pm_J Z$, it follows by replacing Z by $\pm_J Z^H$ in (2.8) and (2.10) that (2.9) and (2.10) define a $n \times n$ minorant of Z. $\square$

Lemma 2.3.[16]. Assume $Z \in \mathbb{C}^{n \times n}$ and let $\check{Z}$ be an $n \times n$ minorant of Z. If in addition $\check{Z}$ is an M-matrix, then

$$|Z^{-1}|_M \leq \leq \check{Z}^{-1}. \tag{2.12}$$

The next lemma is an immediate consequence of Lemmas 2.2 and 2.3.

Lemma 2.4. Assume $Z \in \mathbb{C}^{n \times n}$ and $\check{Z} \in \mathbb{R}^{n \times n}$ satisfies

$$\check{z}_{ii} \leq \max\left\{\frac{1}{2}(z_{ii} + z_{ii}^*), \quad \frac{1}{2}|_J(z_{ii} - z_{ii}^*)|\right\}, \tag{2.13a}$$

$$\check{z}_{ij} \leq -|z_{ij}|, \quad i \neq j. \tag{2.13b}$$

Then, $\check{Z}$ is a $n \times n$ minorant of Z. Furthermore, if $\check{Z}$ is an M-matrix, then

$$|Z^{-1}|_M \leq \leq \check{Z}^{-1}. \tag{2.14}$$

Lemma 2.5. let $Q \in \mathbb{C}^{n \times n}$ and let $\underline{q}$ be a positive scalar satisfying either

$$\underline{q} \leq \frac{1}{2}\lambda_{\min}[Q + Q^H], \tag{2.15}$$

or

$$\underline{q} \leq \frac{1}{2}\lambda_{\min}\left[\pm_J(Q - Q^H)\right]. \tag{2.16}$$

Then,

$$\|Q^{-1}\|_s \leq \underline{q}^{-1}. \tag{2.17}$$

Proof. It follows from Proposition 1 and equation (3.1) of [16] that any positive scalar satisfying

$$\underline{q} \leq -\tau(-Q) \tag{2.18}$$

also satisfies (2.17). Substituting (2.7) into (2.18) yields (2.15). Since $\|\pm_J Q^{-1}\|_s = \|Q^{-1}\|_s$, Q in (2.15) can be replaced by $\pm_J Q$ which yields (2.16). $\square$

5

An immediate extension to Lemma 2.5 is as follows.

Lemma 2.6. Let $Q \in \mathbb{C}^{n \times n}$ and let $\underline{q}$ be a positive scalar satisfying

$$\underline{q} \leq \max\left\{ \frac{1}{2}\lambda_{\min}[Q + Q^H], \frac{1}{2}\lambda_{\min}\left[\jmath(Q - Q^H)\right], \frac{1}{2}\lambda_{\min}\left[\jmath(Q^H - Q)\right]\right\}. \tag{2.19}$$

Then,

$$\|Q^{-1}\|_s \leq \underline{q}^{-1}. \tag{2.19}$$

Lemma 2.7.[23]. Let $A, B \in \mathbb{C}^{n \times n}$. Then,

$$\sigma_{\min}(A + B) \geq \sigma_{\min}(A) - \sigma_{\max}(B). \tag{2.20}$$

The next lemma is a direct consequence of Theorem 4.3.1 of [23].

Lemma 2.8. Let $A, B \in \mathbb{C}^{n \times n}$ be Hermetian and let $\lambda_{\min}(A)$, $\lambda_{\min}(B)$, and $\lambda_{\min}(A + B)$ denote the minimum eigenvalues of the respective arguments. Then,

$$\lambda_{\min}(A + B) \geq \lambda_{\min}(A) + \lambda_{\min}(B). \tag{2.21}$$

Finally, we establish certain definitions and a key lemma used later in the paper. Specifically, a *real-rational matrix fucntion* is a matrix whose elements are rational functions with real coefficients. Furthermore, a *transfer function* is a real-rational matrix each of whose elements is *proper*, i.e., finite at $s = \infty$. A *strictly proper transfer function* is a transfer function that is zero at infinity. An *asymptotically stable transfer function* is a transfer function each of whose poles is in the open left half plane. Finally, a *stable transfer function* is a transfer function each of whose poles is in the closed-left half plane with semi-simple poles on the $\jmath\omega$ axis. Let

$$G(s) \sim \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

denote a state space realization of a transfer function $G(s)$, that is, $G(s) = C(sI - A)^{-1}B + D$. The notation " $\underset{\sim}{\min}$ " is used to denote a minimal realization. The H_2 norm of an assymptotically stable transfer function $G(s)$ is defined as

$$\|G(s)\|_2 \triangleq \left(\frac{1}{2\pi}\int_{-\infty}^{\infty}\|G(\jmath\omega)\|_F^2 \, d\omega\right)^{\frac{1}{2}}. \tag{2.22}$$

6

A square transfer function $G(s)$ is called *positive real* [24] if 1) $G(s)$ is stable, and 2) $G(s) + G^H(s)$ is nonnegative definite for all $\text{Re}[s] > 0$. A square transfer function $G(s)$ is called *strictly positive real* [25-27] if 1) $G(s)$ is asymptotically stable, and 2) $G(\jmath\omega) + G^H(\jmath\omega)$ is positive definite for all real ω. Recall that a minimal realization of a positive real transfer function is stable in the sense of Lyapunov, while a minimal realization of a strictly positive real transfer function is asymptotically stable.

Next we state the well known positive real lemma [28] used to characterize positive realness in the state-space setting.

Lemma 2.9. The strictly proper transfer function $G(s) \overset{\text{min}}{\underset{\sim}{}} \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array} \right]$ is positive real if and only if there exist matrices Q_0 and L with Q_0 positive definite such that

$$AQ_0 + Q_0 A^{\mathrm{T}} = -LL^{\mathrm{T}}, \tag{2.23}$$

$$Q_0 C^{\mathrm{T}} = B. \tag{2.24}$$

This form of the positive real lemma is the dual of that given in [28], and the derivation is similarly dual. See [29] for further details on the dual positive real lemma.

A linear time-invariant system with input w, output y, and transfer function representation

$$y(s) = G(s)w(s) \tag{2.25}$$

is *stable* if $G(s)$ is rational, stable, and proper. This definition of system stability is equivalent to *bounded-input, bounded-output stability*.

3. Positive Real Plants with Strictly Positive Real Feedback

We begin by considering the following n^{th}-order, uncertain, second-order matrix linear plant with proportional damping and rate measurements:

$$\ddot{\eta}(t) + 2\Lambda\Omega\dot{\eta}(t) + \Omega^2\eta(t) = Bu(t) + Dw(t), \tag{3.1a}$$

$$y(t) = C\dot{\eta}(t), \tag{3.1b}$$

$$z(t) = E\dot{\eta}(t), \tag{3.1c}$$

where

$$\Omega = \text{diag}\{\Omega_i\}_{i=1}^n, \quad \Omega_i > 0 \text{ for } i \in \{1, 2, \ldots, n\}, \tag{3.2}$$

$$\Lambda = \text{diag}\{\zeta_i\}_{i=1}^n, \quad \zeta_i > 0 \text{ for } i \in \{1, 2, \ldots, n\}, \tag{3.3}$$

$u \in \mathbb{R}^{n_u}$ is the control vector, $w \in \mathbb{R}^{n_w}$ is the disturbance variable or reference signal, $y \in \mathbb{R}^{n_y}$ represents the rate measurements, and $z \in \mathbb{R}^{n_z}$ represents the performance variables (restricted to be linear functions of the modal rates). It is assumed that

$$\Omega \in \boldsymbol{\Omega} \triangleq \{\Omega_0 + \Delta\Omega : |\Delta\Omega|_\text{M} \leq\leq \widehat{\Delta\Omega}\}, \tag{3.4}$$

$$\Lambda \in \boldsymbol{\Lambda} \triangleq \{\Lambda_0 + \Delta\Lambda : |\Delta\Lambda|_\text{M} \leq\leq \widehat{\Delta\Lambda}\}, \tag{3.5}$$

$$B \in \mathbf{B} \triangleq \{B_0 + \Delta B : |\Delta B|_\text{M} \leq\leq \widehat{\Delta B}\}, \tag{3.6}$$

$$D \in \mathbf{D} \triangleq \{D_0 + \Delta D : |\Delta D|_\text{M} \leq\leq \widehat{\Delta D}\}, \tag{3.7}$$

$$C \in \mathbf{C} \triangleq \{C_0 + \Delta C : |\Delta C|_\text{M} \leq\leq \widehat{\Delta C}\}, \tag{3.8}$$

$$E \in \mathbf{E} \triangleq \{E_0 + \Delta E : |\Delta E|_\text{M} \leq\leq \widehat{\Delta E}\}. \tag{3.9}$$

Next, define

$$H_1 \triangleq (\Omega, \Lambda), \tag{3.10}$$

$$H_2 \triangleq (B, C), \tag{3.11}$$

$$H_3 \triangleq (D, E), \tag{3.12}$$

and define $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_3$ to be the corresponding uncertainty sets, i.e.,

$$\mathbf{H}_1 \triangleq \{(\Omega, \Lambda) : \Omega \in \boldsymbol{\Omega}, \ \Lambda \in \boldsymbol{\Lambda}\}, \tag{3.13}$$

$$\mathbf{H}_2 \triangleq \{(B, C) : B \in \mathbf{B}, \ C \in \mathbf{C}\}, \tag{3.14}$$

$$\mathbf{H}_3 \triangleq \{(D, E) : D \in \mathbf{D}, \ E \in \mathbf{E}\}. \tag{3.15}$$

Additionally, define

$$\mathbf{H} \triangleq \mathbf{H}_1 \cup \mathbf{H}_2 \cup \mathbf{H}_3. \tag{3.16}$$

Note that $\mathbf{H}_1$ is the uncertainty set corresponding to errors in the frequencies and damping ratios while $\mathbf{H}_2$ and $\mathbf{H}_3$ are uncertainty sets corresponding to errors in the mode shapes. It follows from (3.4)–(3.9) that $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_3$ are arcwise connnected.

8

Furthermore, define

$$\theta(s) \triangleq \mathcal{L}[\dot{\eta}(t)], \tag{3.17}$$

so that (3.1) has the s-domain representation

$$\Phi^{-1}(H_1, s)\theta(H, s) = Bu(s) + Dw(s), \tag{3.18a}$$

$$y(H, s) = C\theta(s), \tag{3.18b}$$

$$z(H, s) = E\theta(s), \tag{3.18c}$$

where

$$\Phi(H_1, s) \triangleq \text{diag}\{\phi_i(H_1, s)\}_{i=1}^n, \tag{3.19}$$

and

$$\phi_i(H_1, s) \triangleq \frac{s}{s^2 + 2\zeta_i\Omega_i s + \Omega_i^2}. \tag{3.20}$$

Note that for all $H_1 \in \mathbf{H}_1$, $\Phi(H_1, s)$ is strictly positive real, so that

$$\Phi(H_1, j\omega) + \Phi^H(H_1, j\omega) > 0, \quad H_1 \in \mathbf{H}_1, \quad \omega \in (0, \infty). \tag{3.21}$$

If, alternatively, the system is undamped, that is, $\zeta_i = 0$, $i = 1, \ldots, n$, then (3.19) is positive real.

To make the model more realistic we now include sensor and actuator dynamics that are assumed to be known. (These dynamics could be empirically determined via hardware experimentation.) The matrix of actuator dynamics (Ψ_a) and the matrix of sensor dynamics (Ψ_s) are given respectively by

$$\Psi_a(s) \triangleq \text{diag}\{\Psi_{a,i}(s)\}_{i=1}^{n_u}, \tag{3.22}$$

$$\Psi_s(s) \triangleq \text{diag}\{\Psi_{s,i}(s)\}_{i=1}^{n_y}. \tag{3.23}$$

Appending these dynamics to the system (3.13) yields

$$\Phi^{-1}(H_1, s)\theta(H, s) = B\Psi_a(s)u(s) + Dw(s), \tag{3.24a}$$

$$y(H, s) = \Psi_s(s)C\theta(s), \tag{3.24b}$$

$$z(H, s) = E\theta(s). \tag{3.24c}$$

9

Next, assume that the linear feedback law

$$u(s) = -K(s)y(s) \tag{3.25}$$

stabilizes the nominal system, i.e., the system (3.24) with $H_1 = (\Omega_0, \Lambda_0)$ and $H_2 = (B_0, C_0)$. Substituting (3.25) into (3.24a) gives

$$[\Phi^{-1}(H_1, s) + F(H_2, s)]\theta(H, s) = Dw(s), \tag{3.26}$$

where

$$F(H_2, s) \triangleq B\Psi_a(s)K(s)\Psi_s(s)C. \tag{3.27}$$

Now define $G_{w\theta}(H, s)$ to be the transfer function between $w(s)$ and $\theta(H, s)$, such that

$$\theta(H, s) = G_{w\theta}(H, s)w(s). \tag{3.28}$$

Then, the following proposition is needed for Theorem 3.1.

Proposition 3.1. For given $H \in \mathbf{H}$, $G_{w\theta}(H, s)$ is asymptotically stable if

$$\det\left[\Phi^{-1}(H_1, \jmath\omega) + F(H_2, \jmath\omega)\right] \neq 0, \quad \omega \in [0, \infty). \tag{3.29}$$

Proof. The result is a direct consequence of the multivariable Nyquist criterion. $\square$

The proof of the following theorem relies on Proposition 3.1. For the statement of the next result let $Q : \mathbf{H} \to \mathrm{IR}^{n \times n}$, and define $\nu(Q)$ by

$$\nu(Q) \triangleq \max\left\{ \min_{H \in \mathbf{H}} \frac{1}{2}\lambda_{\min}\left(Q(H) + Q^H(H)\right), \quad \min_{H \in \mathbf{H}} \frac{1}{2}\lambda_{\min}\left(\jmath(Q(H) - Q^H(H))\right), \right.$$
$$\left. \min_{H \in \mathbf{H}} \frac{1}{2}\lambda_{\min}\left(\jmath(Q^H(H) - Q(H))\right)\right\}. \tag{3.30}$$

Theorem 3.1. If for all $H_2 \in \mathbf{H}_2$, $F(H_2, s)$ is positive real, then $G_{w\theta}(H, s)$ is asymptotically stable for all $H \in \mathbf{H}$ (that is $G_{w\theta}(H, s)$ is *robustly stable*). In addition,

$$\left\|\left[\Phi^{-1}(H_1, \jmath\omega) + F(H_2, \jmath\omega)\right]^{-1}\right\|_s \leq \underline{p}^{-1}(\jmath\omega), \tag{3.31}$$

10

where $\underline{p}(\jmath\omega)$ is any positive scalar satisfying

$$\underline{p}(\jmath\omega) \leq \nu[\Phi^{-1}(\jmath\omega) + F(\jmath\omega)]. \tag{3.32}$$

Proof. First note that $\Phi(H_1, s)$ is strictly positive real for all $H_1 \in \mathbf{H}_1$. We now show that $\Phi(H_1, \jmath\omega)$ is invertible and $\Phi^{-1}(H_1, s)$ is strictly positive real.

Let $H_1 \in \mathbf{H}_1$, $x \in \mathbb{C}^n$, $x \neq 0$, and $\lambda \in \mathbb{C}$ be such that $\Phi(H_1, \jmath\omega)x = \lambda x$ and hence $x^H\Phi(H, \jmath\omega) = \lambda^H x^H$. Then, $x^H[\Phi(H_1, \jmath\omega) + \Phi^H(H_1, \jmath\omega)]x > 0$ implies Re $\lambda > 0$. Hence $\det[\Phi(H_1, \jmath\omega)] \neq 0$. In addition,

$$\Phi^{-1}(H_1, \jmath\omega) + \Phi^{-H}(H_1, \jmath\omega) = \Phi^{-1}(H_1, \jmath\omega)[\Phi(H_1, \jmath\omega) + \Phi^H(H_1, \jmath\omega)]\Phi^{-H}(H_1, \jmath\omega) > 0,$$

which implies that $\Phi^{-1}(H_1, s)$ is strictly positive real. Since, for all $H_2 \in \mathbf{H}_2$, $F(H_2, s)$ is positive real it follows that for all $H_1 \in \mathbf{H}_1$ and $H_2 \in \mathbf{H}_2$, $[\Phi^{-1}(H_1, s) + F(H_2, s)]$ is strictly positive real and thus (3.29) is satisfied. Hence, it follows from Proposition 3.1 that $G_{w\theta}(s)$ is asymptotically stable for all $H \in \mathbf{H}$.

Now, for $H_1 \in \mathbf{H}_1$ and $H_2 \in \mathbf{H}_2$, $\Phi^{-1}(H_1, s) + F(H_2, s)$ is strictly positive real and hence $\nu[\Phi^{-1}(\jmath\omega) + F(\jmath\omega)]$ is positive. Equation (3.31) then follows using Lemma 2.6. $\square$

Remark 3.1. The first part of the proof to Theorem 3.1 is essentially identical to the proof of Lemma 3.2 of [30].

Remark 3.2. The norm bound (3.31) lays the foundation for one of the performance bounds given in the next section.

Remark 3.3. Note that Theorem 3.1 also holds if, alternatively, the plant is positive real and the compensator is strictly positive real.

Remark 3.4. Note that in the scalar case the definition of $\nu(Q)$ can be specialized to

$$\nu(Q) \triangleq \max\left\{\min_{H \in \mathbf{H}} \frac{1}{2}\left(Q(H) + Q^*(H)\right), \min_{H \in \mathbf{H}} \frac{1}{2}\left|\jmath(Q(H) - Q^*(H))\right|\right\}. \tag{3.33}$$

4. Performance Bounds

In this section it is again assumed that $F(H_2, s)$ is positive real for all $H_2 \in \mathbf{H_2}$. We define

$$\Gamma(H_1, H_2, s) \triangleq \left[\Phi^{-1}(H_1, s) + F(H_2, s)\right]^{-1}, \tag{4.1}$$

and note that in this case (3.26) yields

$$\theta(H, s) = \Gamma(H_1, H_2, s) D w(s). \tag{4.2}$$

Let $\hat{\Gamma}(\jmath\omega)$ denote a majorant of $\Gamma(H_1, H_2, \jmath\omega)$ for all $H_1 \in \mathbf{H_1}$ and $H_2 \in \mathbf{H_2}$, such that

$$\max_{\substack{H_1 \in \mathbf{H_1} \\ H_2 \in \mathbf{H_2}}} \mid \Gamma(H_1, H_2, \jmath\omega) \mid_{\mathrm{M}} \leq \leq \hat{\Gamma}(\jmath\omega). \tag{4.3}$$

Then, applying the inequality (2.2) to (4.2) gives

$$\max_{H \in \mathbf{H}} \mid \theta(H, \jmath\omega) \mid_{\mathrm{M}} \leq \leq \hat{\Gamma}(\jmath\omega)\hat{D} \mid w(\jmath\omega)\mid_{\mathrm{M}}. \tag{4.4}$$

Similarly, applying the inequality (2.2) to (3.24c) and using (4.4) gives

$$\mid z(\jmath\omega) \mid_{\mathrm{M}} \leq \leq \mid E\mid_{\mathrm{M}}\hat{\theta}(\jmath\omega), \tag{4.5}$$

where

$$\hat{\theta}(\jmath\omega) \triangleq \hat{\Gamma}(\jmath\omega)\hat{D}\mid w(\jmath\omega)\mid_{\mathrm{M}}. \tag{4.6}$$

Equations (4.5) and (4.6) indicate that performance bounding requires the computation of $\hat{\Gamma}(\jmath\omega)$ satisfying (4.3). The following two theorems present alternatives for the choice of $\hat{\Gamma}(\jmath\omega)$. The first theorem follows directly from Theorem 3.1.

Theorem 4.1. Assume that for all $H_2 \in \mathbf{H_2}$, $F(H_2, s)$ is positive real and $\underline{p}(\jmath\omega)$ is a positive scalar satisfying (3.32). Then

$$\max_{\substack{H_1 \in \mathbf{H_1} \\ H_2 \in \mathbf{H_2}}} \mid \Gamma(H_1, H_2, \jmath\omega) \mid_{\mathrm{M}} \leq \leq \hat{\Gamma}_0(\jmath\omega), \tag{4.7}$$

where

$$\hat{\Gamma}_0(\jmath\omega) = \underline{p}^{-1}(\jmath\omega)U_n, \tag{4.8}$$

12

and U_n denotes the $n \times n$ matrix with all unity elements.

Let $\mathbb{D}^n$ denote the set of $n \times n$ diagonal matrices, let $(\cdot)$ have the mapping $Q : \mathbf{H} \to \mathbb{D}^n$, and define the function $\nu_D(Q)$ by

$$\nu_D(Q) \triangleq \operatorname{diag}\{\nu(q_{ii})\}_{i=1}^{n}. \tag{4.9}$$

We are now prepared to state the next theorem.

Theorem 4.2. Assume that for all $H_2 \in \mathbf{H_2}$, $F(H_2, s)$ is positive real and let $F_\mathrm{d}(H_2, s)$, and $F_\mathrm{od}(H_2, s)$ respectively denote the diagonal and off-diagonal matrices corresponding to $F(H_2, s)$, such that

$$F_\mathrm{d}(H_2, s) \triangleq \operatorname{diag}\{f_{ii}(H_2, s)\}_{i=1}^{n}, \tag{4.10}$$

$$F_\mathrm{od}(H_2, s) \triangleq F(H_2, s) - F_\mathrm{d}(H_2, s). \tag{4.11}$$

Let $\Pi(\jmath\omega)$ be given by

$$\Pi(\jmath\omega) = P(\jmath\omega) - \hat{F}_\mathrm{od}(\jmath\omega), \tag{4.12}$$

where $P(\jmath\omega)$ is diagonal and satisfies

$$P(\jmath\omega) \leq\leq \nu_D\big(\Phi^{-1}(H_1, \jmath\omega) + F_\mathrm{d}(H_2, \jmath\omega)\big), \tag{4.13}$$

and $\hat{F}_\mathrm{od}(\jmath\omega)$ satisfies

$$[\hat{F}_\mathrm{od}(\jmath\omega)]_{ij} \geq \max_{H_2 \in \mathbf{H_2}} \big|[F_\mathrm{od}(H_2, \jmath\omega)]_{ij}\big|. \tag{4.14}$$

Then,

$$\max_{\substack{H_1 \in \mathbf{H_1} \\ H_2 \in \mathbf{H_2}}} |\Gamma(\jmath\omega)|_\mathrm{M} \leq\leq \Pi^{-1}(\jmath\omega). \tag{4.15}$$

Proof. First note that $\Gamma(H_1, H_2, s)$ is given by

$$\Gamma(H_1, H_2, s) = \big[S_\mathrm{d}(H_1, H_2, s) + F_\mathrm{od}(H_2, s)\big]^{-1}, \tag{4.16}$$

where

$$S_\mathrm{d}(H_1, H_2, s) \triangleq \Phi^{-1}(H_1, s) + F_\mathrm{d}(H_2, s). \tag{4.17}$$

13

Let $\check{S}_{\mathrm{d}}(\jmath\omega)$ be a minorant of $S_{\mathrm{d}}(H_1, H_2, \jmath\omega)$, and let $\hat{F}_{\mathrm{od}}(\jmath\omega)$ satisfy (4.14) for all $H \in \mathbf{H}$. Furthermore, if $\check{S}_{\mathrm{d}}(\jmath\omega) - \hat{F}_{\mathrm{od}}(\jmath\omega)$ is an M-matrix it follows from Lemma 2.3 that

$$|\Gamma(\jmath\omega)|_{\mathrm{M}} \leq\leq [\check{S}_{\mathrm{d}}(\jmath\omega) - \hat{F}_{\mathrm{od}}(\jmath\omega)]^{-1}. \tag{4.18}$$

Since $\nu_D(S_{\mathrm{d}}(\jmath\omega))$ is a minorant of $S_{\mathrm{d}}(H_1, H_2, \jmath\omega)$, so is $P(\jmath\omega)$ and hence the proof is complete. $\square$

Remark 4.1. Note that for the case $n = 1$ Theorems 4.1 and 4.2 yield the same bound. However, in the case $n > 1$ the performance bound is obtained by computing the minimum of the bounds given by Theorems 4.1 and 4.2.

5. Performance Bounds for Colocated Rate Feedback

In this section we give performance bounds for decentralized colocated rate feedback systems. Specifically we assume that

$$C = B^{\mathrm{T}}, \tag{5.1}$$

$$\Psi_a(s) = \Psi_s(s) = I_n, \tag{5.2}$$

$$K(s) = \kappa\kappa^{\mathrm{T}}, \tag{5.3}$$

where

$$\kappa = \mathrm{diag}\{\kappa_i\}_{i=1}^m. \tag{5.4}$$

Hence,

$$F(H_2, s) = B\kappa\kappa^{\mathrm{T}}B^{\mathrm{T}}. \tag{5.5}$$

We now show how to practically compute bounds corresponding to Theorems 4.1 and 4.2. First, however, define $\mathcal{S} : \mathrm{IR} \to \mathrm{IR}$ as

$$\mathcal{S}(\alpha) \triangleq \begin{cases} \alpha, & \alpha \geq 0 \\ 0, & \alpha < 0. \end{cases} \tag{5.6}$$

We begin by showing how to compute a positive scalar $\underline{p}(\jmath\omega)$ satisfying (3.32) with $\nu(\cdot)$ given by (3.30). Using Lemma 2.8 it follows that

$$\begin{aligned} &\lambda_{\min}[\Phi^{-1}(H_1, \jmath\omega) + F(H_2, \jmath\omega) + \Phi^{-H}(H_1, \jmath\omega) + F^H(H_2, \jmath\omega)] \\ &\geq \lambda_{\min}[\Phi^{-1}(H_1, \jmath\omega) + \Phi^{-H}(H_1, \jmath\omega)] + \lambda_{\min}[F(H_2, \jmath\omega) + F^H(H_2, \jmath\omega)]. \end{aligned} \tag{5.7}$$

Now,

$$\Phi^{-1}(H_1, \jmath\omega) + \Phi^{-H}(H_1, \jmath\omega) = \mathrm{diag}\{4\zeta_k\Omega_k\}_{k=1}^n$$
$$= \mathrm{diag}\{4(\zeta_{0,k} + \Delta\zeta_{0,k})(\Omega_{0,k} + \Delta\Omega_{0,k})\}. \tag{5.8}$$

Hence,

$$\min_{H_1 \in \mathbf{H}_1} \frac{1}{2}\lambda_{\min}[\Phi^{-1}(H_1, \jmath\omega) + \Phi^{-H}(H_1, \jmath\omega)] = \min_k 2(\zeta_{0,k} - \widehat{\Delta\zeta}_k)(\Omega_{0,k} - \widehat{\Delta\Omega}_k). \tag{5.9}$$

Furthermore,

$$F(H_2, \jmath\omega) + F^H(H_2, \jmath\omega) = 2B\kappa\kappa^{\mathrm{T}}B^{\mathrm{T}}. \tag{5.10}$$

Note that

$$\lambda_{\min}(2B\kappa\kappa^{\mathrm{T}}B^{\mathrm{T}}) = 2\sigma_{\min}^2(B\kappa) = 2\sigma_{\min}^2\big((B_0 + \Delta B)\kappa\big). \tag{5.11}$$

Now, using (2.20) it follows that

$$\sigma_{\min}\big((B_0 + \Delta B)\kappa\big) \geq \sigma_{\min}(B_0\kappa) - \sigma_{\max}(\Delta B\kappa)$$
$$\geq \sigma_{\min}(B_0\kappa) - (\max_i \kappa_i)\sigma_{\max}(\Delta B) \tag{5.12}$$
$$\geq \sigma_{\min}(B_0\kappa) - (\max_i \kappa_i)\|\widehat{\Delta B}\|_{\mathrm{F}}.$$

Since $\sigma_{\min}\big((B_0 + \Delta B)\kappa\big) > 0$, (5.11), (5.12), and (5.6) yield

$$\min_{H_2 \in \mathbf{H}_2} \frac{1}{2}\lambda_{\min}\big[F(H_2, \jmath\omega) + F^H(H_2, \jmath\omega)\big] \geq \Big[\mathcal{S}\big(\sigma_{\min}(B_0\kappa) - (\max_i \kappa_i)\|\widehat{\Delta B}\|_{\mathrm{F}}\big)\Big]^2. \tag{5.13}$$

It now follows from (5.6), (5.8), and (5.12) that

$$\min_{H \in \mathbf{H}} \frac{1}{2}\lambda_{\min}\big[\Phi^{-1}(H_1, \jmath\omega) + F(H_2, \jmath\omega) + \Phi^{-H}(H_1, \jmath\omega) + F^H(H_2, \jmath\omega)\big]$$
$$\geq \min_k 2(\zeta_{0,k} - \widehat{\Delta\zeta}_k)(\Omega_{0,k} - \widehat{\Delta\Omega}_k) \tag{5.14}$$
$$+ \Big[\mathcal{S}\big(\sigma_{\min}(B_0\kappa) - (\max_i \kappa_i)\|\widehat{\Delta B}\|_{\mathrm{F}}\big)\Big]^2.$$

Once again using Lemma 2.8, it follows that

$$\lambda_{\min}\big[\jmath\Phi^{-1}(H_1, \jmath\omega) - \jmath\Phi^{-H}(H_1, \jmath\omega) + \jmath F(H_2, \jmath\omega) - \jmath F^H((H_2, \jmath\omega)\big]$$
$$\geq \lambda_{\min}\big[\jmath\Phi^{-1}(H_1, \jmath\omega) - \jmath\Phi^{-H}(H_1, \jmath\omega)\big] + \lambda_{\min}\big[\jmath F(H_2, \jmath\omega) - \jmath F^H(H_2, \jmath\omega)\big]. \tag{5.15}$$

Now, noting

$$\jmath\Phi^{-1}(H_1, \jmath\omega) - \jmath\Phi^{-H}(H_1, \jmath\omega) = 2\,\mathrm{diag}\{\frac{1}{\omega}\Omega_k^2 - \omega\}_{k=1}^n, \tag{5.16}$$

15

we obtain

$$\min_{H_1 \in \mathbf{H_1}} \frac{1}{2} \lambda_{\min} \left[\jmath \Phi^{-1}(H_1, \jmath\omega) - \jmath \Phi^{-H}(H_1, \jmath\omega) \right] = \min_k \frac{1}{\omega}(\Omega_k - \widehat{\Delta\Omega_k})^2 - \omega. \qquad (5.17)$$

Furthermore,

$$\jmath F(H_2, \jmath\omega) - \jmath F^H(H_2, \jmath\omega) = 0, \qquad (5.18)$$

yields

$$\min_{H_2 \in \mathbf{H_2}} \frac{1}{2} \lambda_{\min} \left[\jmath F(H_2, \jmath\omega) - \jmath F^H(H_2, \jmath\omega) \right] = 0. \qquad (5.19)$$

It now follows from (5.15), (5.17), and (5.19) that

$$\min_{H \in \mathbf{H}} \frac{1}{2} \lambda_{\min} \left[\jmath \Phi^{-1}(H_1, \jmath\omega) - \jmath \Phi^{-H}(H_1, \jmath\omega) + \jmath F(H_2, \jmath\omega) - \jmath F^H(H_2, \jmath\omega) \right]$$
$$\geq \min_k \frac{1}{\omega}(\Omega_k - \widehat{\Delta\Omega_k})^2 - \omega. \qquad (5.20)$$

Similarly,

$$\min_{H \in \mathbf{H}} \frac{1}{2} \lambda_{\min} \left[\jmath \Phi^{-H}(H_1, \jmath\omega) - \jmath \Phi^{-1}(H_1, \jmath\omega) + \jmath F^H(H_2, \jmath\omega) - \jmath F(H_2, \jmath\omega) \right]$$
$$\geq \min_k \omega - \frac{1}{\omega}(\Omega_k + \widehat{\Delta\Omega_k})^2. \qquad (5.21)$$

The following theorem is now immediate using (3.30) with (5.14), (5.20), and (5.21).

Theorem 5.1. If $F(H_2, s)$ is given by (5.5) then $\underline{p}(\jmath\omega)$ satisfying Theorems 3.1 and 4.1 is given by

$$\underline{p}(\jmath\omega) = \max \Big\{ \min_k 2(\zeta_{0,k} - \widehat{\Delta\zeta_k})(\Omega_{0,k} - \widehat{\Delta\Omega_k})$$
$$+ \left[\mathcal{S}(\sigma_{\min}(B_0\kappa) - (\max_i \kappa)\|\widehat{\Delta B}\|_{\mathrm{F}}) \right]^2, \qquad (5.22)$$
$$\min_k \frac{1}{\omega}(\Omega_k - \widehat{\Delta\Omega_k})^2 - \omega, \quad \min_k \omega - \frac{1}{\omega}(\Omega_k + \widehat{\Delta\Omega_k})^2 \Big\},$$

and $F_{\mathrm{d}}(H_2, \jmath\omega)$ is given by

$$F_{\mathrm{d}}(H_2, \jmath\omega) = \mathrm{diag} \Big\{ \sum_{j=1}^m b_{ij} \kappa_j^2 b_{ij} \Big\}_{i=1}^n. \qquad (5.23)$$

16

Furthermore, the (i, j) element of $F_{\mathrm{od}}(H_2, \jmath\omega)$ is given by

$$\left[F_{\mathrm{od}}(H_2, \jmath\omega) \right]_{ij} = \sum_{l=1}^{m} b_{i\ell} \kappa_\ell^2 b_{j\ell}. \tag{5.24}$$

Using (5.23) and (5.24) and following a similar procedure used to develop Theorem 5.1, we obtain the following result.

Theorem 5.2. Let $F(H_2, s)$ be given by (5.5) then $P(\jmath\omega)$ satisfying Theorem 4.2 is given by

$$P(\jmath\omega) = \mathrm{diag}\{ p_{kk}(\jmath\omega) \}_{k=1}^{n}, \tag{5.25}$$

where

$$p_{kk}(\jmath\omega) = \max\Bigg\{ 2(\zeta_{0,k} - \widehat{\Delta\zeta}_k)(\Omega_{0,k} - \widehat{\Delta\Omega}_k) \\ + \sum_{j=1}^{m} \left[\mathcal{S}\big((b_{kj} - \widehat{\Delta B}_{kj})\kappa_j \big) \right]^2, \\ \min_{\Omega \in \Omega} \left| \frac{1}{\omega} \Omega_k^2 - \omega \right| \Bigg\}. \tag{5.26}$$

In addition, the (i, j) element of $\hat{F}_{\mathrm{od}}(\jmath\omega)$ satisfying (4.14) of Theorem 4.2 is given by

$$\left[\hat{F}_{\mathrm{od}}(\jmath\omega) \right]_{ij} = \sum_{\ell=1}^{m} (|B_{0,i\ell}| + \widehat{\Delta B}_{i\ell}) \kappa_\ell^2 (|B_{0,j\ell}| + \widehat{\Delta B}_{j\ell}). \tag{5.27}$$

6. Extensions to Dynamic Compensation

In this section we generalize the results of Section 5 to dynamic compensation. Once again we assume colocated rate feedback with negligible sensor and actuator dynamics so that (5.1) and (5.2) hold. The following two theorems provide performance bounds for positive real systems controlled by strictly positive real dynamic compensators.

Theorem 6.1. If $F(H_2, s) = BK(s)B^{\mathrm{T}}$ then $\underline{p}(\jmath\omega)$ satisfying Theorems 3.1 and 4.1

17

is given by

$$
\underline{p}(\jmath\omega) = \max\Bigg\{ \min_k 2(\zeta_{0,k} - \widehat{\Delta\zeta}_k)(\Omega_{0,k} - \widehat{\Delta\Omega}_k)
$$

$$
+ \frac{1}{2}\Big[\mathcal{S}\big(\sigma_{\min}(B_0 M(\jmath\omega)) - \sigma_{\max}(M(\jmath\omega))\|\widehat{\Delta B}\|_{\mathrm{F}}\big)\Big]^2,
$$

$$
\min_k(\frac{1}{\omega}(\Omega_{0,k} - \widehat{\Delta\Omega}_k)^2 - \omega) - \frac{1}{2}\sigma_{\max}(K(\jmath\omega) - K^H(\jmath\omega))(\sigma_{\max}(B_0) + \|\widehat{\Delta B}\|_{\mathrm{F}})^2,
$$

$$
\min_k(\omega - \frac{1}{\omega}(\Omega_{0,k} + \widehat{\Delta\Omega}_k)^2) - \frac{1}{2}\sigma_{\max}(K(\jmath\omega) - K^H(\jmath\omega))(\sigma_{\max}(B_0) + \|\widehat{\Delta B}\|_{\mathrm{F}})^2\Bigg\},
\tag{6.1}
$$

where

$$
K(\jmath\omega) + K^H(\jmath\omega) = M(\jmath\omega)M^H(\jmath\omega).
\tag{6.2}
$$

Proof. From (5.7) and (5.9) we have

$$
\min_{H\in\mathbf{H}} \frac{1}{2}\lambda_{\min}[\Phi^{-1}(H_1,\jmath\omega) + F(H_2,\jmath\omega) + \Phi^{-H}(H_1,\jmath\omega) + F^H(H_2,\jmath\omega)]
$$

$$
\geq \min_k 2(\zeta_{0,k} - \widehat{\Delta\zeta}_k)(\Omega_{0,k} - \widehat{\Delta\Omega}_k)
\tag{6.3}
$$

$$
+ \min_{H_2\in\mathbf{H}_2} \frac{1}{2}\lambda_{\min}[F(H_2,\jmath\omega) + F^H(H_2,\jmath\omega)].
$$

Next, note that

$$
\begin{aligned}
\lambda_{\min}[F(H_2,\jmath\omega) + F^H(H_2,\jmath\omega)] &= \lambda_{\min}\big(B(K(\jmath\omega) + K^H(\jmath\omega))B^{\mathrm{T}}\big), \\
&= \lambda_{\min}\big(BM(\jmath\omega)M^H(\jmath\omega)B^{\mathrm{T}}\big), \\
&= \sigma_{\min}^2(BM(\jmath\omega)), \\
&= \sigma_{\min}^2((B_0 + \Delta B)M(\jmath\omega)).
\end{aligned}
\tag{6.4}
$$

Now, using (2.20) it follows that

$$
\begin{aligned}
\sigma_{\min}((B_0 + \Delta B)M) &\geq \sigma_{\min}(B_0 M(\jmath\omega)) - \sigma_{\max}(\Delta B M(\jmath\omega)), \\
&\geq \sigma_{\min}(B_0 M(\jmath\omega)) - \sigma_{\max}(M(\jmath\omega))\sigma_{\max}(\Delta B), \\
&\geq \sigma_{\min}(B_0 M(\jmath\omega)) - \sigma_{\max}(M(\jmath\omega))\|\widehat{\Delta B}\|_{\mathrm{F}}.
\end{aligned}
\tag{6.5}
$$

Noting that $\sigma_{\min}(BM(\jmath\omega)) \geq 0$, (6.4), (6.5), and (5.6) yield

$$
\min_{H_2\in\mathbf{H}_2} \frac{1}{2}\lambda_{\min}[F(H_2,\jmath\omega) + F^H(H_2,\jmath\omega)]
$$

$$
\geq \frac{1}{2}\Big[\mathcal{S}\big(\sigma_{\min}(B_0 M(\jmath\omega)) - \sigma_{\max}(M(\jmath\omega))\|\widehat{\Delta B}\|_{\mathrm{F}}\big)\Big]^2.
\tag{6.6}
$$

18

It now follows from (6.3) and (6.6) that

$$\min_{H \in \mathbf{H}} \frac{1}{2} \lambda_{\min}[\Phi^{-1}(H_1, \jmath\omega) + F(H_2, \jmath\omega) + \Phi^{-H}(H_1, \jmath\omega) + F^H(H_2, \jmath\omega)]$$

$$\geq \min_k 2(\zeta_{0,k} - \widehat{\Delta\zeta}_k)(\Omega_{0,k} - \widehat{\Delta\Omega}_k) \tag{6.7}$$

$$+ \frac{1}{2}\left[\mathcal{S}\big(\sigma_{\min}(B_0 M(\jmath\omega)) - \sigma_{\max}(M(\jmath\omega))\|\widehat{\Delta B}\|_{\mathrm{F}}\big)\right]^2.$$

Similarly, using (5.15) and (5.17) we obtain

$$\min_{H \in \mathbf{H}} \frac{1}{2} \lambda_{\min}[\jmath\Phi^{-1}(H_1, \jmath\omega) - \jmath\Phi^{-H}(H_1, \jmath\omega) + \jmath F(H_2, \jmath\omega) - \jmath F^H(H_2, \jmath\omega)]$$

$$\geq \min_k \frac{1}{\omega}(\Omega_{0,k} - \widehat{\Delta\Omega}_k)^2 - \omega + \min_{H_2 \in \mathbf{H}_2} \frac{1}{2} \lambda_{\min}[\jmath(F(H_2, \jmath\omega) - F^H(H_2, \jmath\omega))]. \tag{6.8}$$

Hence,

$$\frac{1}{2}\lambda_{\min}[\jmath(F(H_2, \jmath\omega) - F^H(H_2, \jmath\omega))] \geq -\frac{1}{2}\sigma_{\max}(F(H_2, \jmath\omega) - F^H(H_2, \jmath\omega)),$$

$$= -\frac{1}{2}\sigma_{\max}\big(B(K(\jmath\omega) - K^H(\jmath\omega))B^{\mathrm{T}}\big),$$

$$\geq -\frac{1}{2}\sigma_{\max}^2(B)\sigma_{\max}(K(\jmath\omega) - K^H(\jmath\omega)),$$

$$\geq -\frac{1}{2}\big(\sigma_{\max}(B_0) + \|\widehat{\Delta B}\|_{\mathrm{F}}\big)^2 \sigma_{\max}(K(\jmath\omega) - K^H(\jmath\omega)). \tag{6.9}$$

Finally using (6.8) and (6.9) it follows that

$$\min_{H \in \mathbf{H}} \frac{1}{2} \lambda_{\min}[\jmath(\Phi^{-1}(H_1, \jmath\omega) - \Phi^{-H}(H_1, \jmath\omega)) + \jmath(F(H_2, \jmath\omega) - F^H(H_2, \jmath\omega))]$$

$$\geq \min_k \frac{1}{\omega}(\Omega_{0,k} - \widehat{\Delta\Omega}_k)^2 - \omega \tag{6.10}$$

$$- \frac{1}{2}\sigma_{\max}(K(\jmath\omega) - K^H(\jmath\omega))\big(\sigma_{\max}(B_0) + \|\widehat{\Delta B}\|_{\mathrm{F}}\big)^2.$$

Using a similar procedure given above we obtain

$$\min_{H \in \mathbf{H}} \frac{1}{2} \lambda_{\min}[\jmath(\Phi^{-H}(H_1, \jmath\omega) - \Phi^{-1}(H_1, \jmath\omega)) + \jmath(F^H(H_2, \jmath\omega) - F(H_2, \jmath\omega))]$$

$$\geq \min_k \quad \omega - \frac{1}{\omega}(\Omega_{0,k} + \widehat{\Delta\Omega}_k)^2 \tag{6.11}$$

$$- \frac{1}{2}\sigma_{\max}(K(\jmath\omega) - K^H(\jmath\omega))\big(\sigma_{\max}(B_0) + \|\widehat{\Delta B}\|_{\mathrm{F}}\big)^2.$$

Hence, using (6.7), (6.10), and (6.11) and Theorem 3.1 the theorem is proved. $\square$

The next theorem gives the second performance bound for the dynamic compenstor case. Using Theorem 4.2 and a procedure similar to the one employed in the previous theorem, the following result is immediate.

Theorem 6.2. Let $F(H_2, s) = BK(s)B^{\mathrm{T}}$. Then $P(\jmath\omega)$ satisfying Theorem 4.2 is given by

$$P(\jmath\omega) = \mathrm{diag}\{p_{kk}(\jmath\omega)\}_{k=1}^{n}, \tag{6.12}$$

where

$$p_{kk}(\jmath\omega) = \max\Bigg\{2(\zeta_{0,k} - \widehat{\Delta\zeta}_k)(\Omega_{0,k} - \widehat{\Delta\Omega}_k)$$

$$+ \frac{1}{2}\lambda_{\min}(K(\jmath\omega) + K^H(\jmath\omega))\sum_{l=1}^{m}\left[\mathcal{S}(B_{0,kl} - \widehat{\Delta B}_{kl})\right]^2,$$

$$\min_{\Omega\in\Omega}\left|\frac{\Omega_{0,k}^2}{\omega} - \omega\right| - \frac{1}{2}\sigma_{\max}(K(\jmath\omega) - K^H(\jmath\omega))\sum_{l=1}^{m}\left[|B_{0,kl}| + \widehat{\Delta B}_{kl}\right]^2\Bigg\}. \tag{6.13}$$

In addition, $\hat{F}_{\mathrm{od}}(\jmath\omega)$ satisfying (4.14) is given by

$$[\hat{F}_{\mathrm{od}}(\jmath\omega)]_{ij} = \sigma_{\max}(K(\jmath\omega))\Big[\sum_{k=1}^{m}(|B_{0,ik}| + \widehat{\Delta B}_{ik})^2\Big]^{\frac{1}{2}}\Big[\sum_{k=1}^{m}(|B_{0,jk}| + \widehat{\Delta B}_{jk})^2\Big]^{\frac{1}{2}}. \tag{6.14}$$

Remark 6.1. Note that in general the performance bounds given by Theorems 6.1 and 6.2 are more conservative than the bounds given by Theorems 5.1 and 5.2 since they do not exploit the diagonal structure of the controller assumed in Section 5.

Although there is no general theory yet available for designing positive real dynamic compensators, a variety of techniques have been proposed based on H_2 theory [31-38] and H_∞ theory [38-40]. Next, for completeness, we present a systematic approach for designing strictly positive real dynamic compensators for positive real plants. Specifically, for simplicity we restrict our attention to flexible structures with n_u force inputs and n_u velocity measurements so that the colocated admittance, or driving point mobility, is characterized by

$$M\ddot{q}(t) + C\dot{q}(t) + Kq(t) = Bu(t), \tag{6.15}$$

20

$$y(t) = B^{\mathrm{T}} \dot{q}(t), \tag{6.16}$$

where M, C, and K are mass, damping, and stiffness matrices, respectively, and B is determined by the sensor/actuator location. In this case the state space realization of (6.15) and (6.16) is given by

$$G(s) \sim \left[\begin{array}{c|c} \begin{bmatrix} 0 & I \\ -M^{-1}K & -M^{-1}C \end{bmatrix} & \begin{bmatrix} 0 \\ M^{-1}B \end{bmatrix} \\ \hline [0 \quad B^{\mathrm{T}}] & 0 \end{array} \right] = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & 0 \end{array} \right]. \tag{6.17}$$

Note that under the assumption of proportional damping the vibrational model in (6.15) and (6.16) can always be transformed into the form of (3.1). Finally, with $G(s)$ given by (6.17) it follows that (2.23) and (2.24) are satisfied by [35,38]

$$Q_0 = \begin{bmatrix} K^{-1} & 0 \\ 0 & M^{-1} \end{bmatrix}, \qquad L = \begin{bmatrix} 0 \\ \sqrt{2}M^{-1}C^{\frac{1}{2}} \end{bmatrix}. \tag{6.18}$$

Next, we recast (3.1) in a state-space form and address the strict positive real controller synthesis problem. Specifically, given the $2n^{\mathrm{th}}$-order minimal positive real plant

$$\dot{x}(t) = \hat{A}x(t) + \hat{B}u(t) + D_1 w(t), \tag{6.19}$$

$$y(t) = \hat{C}x(t) + D_2 w(t), \tag{6.20}$$

we seek to determine a $2n^{\mathrm{th}}$-order dynamic compensator $-K(s) \sim \left[\begin{array}{c|c} A_c & B_c \\ \hline -C_c & 0 \end{array} \right]$ of the form

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \tag{6.21}$$

$$u(t) = -C_c x_c(t), \tag{6.22}$$

that satisfies the following design criteria:

(i) the closed-loop system (6.19)-(6.22) given by $\tilde{A} \triangleq \begin{bmatrix} \hat{A} & -\hat{B}C_c \\ B_c\hat{C} & A_c \end{bmatrix}$ is asymptotically stable;

(ii) the H$_2$ performance measure

$$J(A_c, B_c, C_c) \triangleq \lim_{t \to \infty} \frac{1}{t} \int_0^t [x^{\mathrm{T}}(s)R_1 x(s) + u^{\mathrm{T}}(s)R_2 u(s)]ds \tag{6.23}$$

21

is minimized, where $R_1 \geq 0$, $R_2 > 0$; and

(iii) $-K(s) \sim \left[\begin{array}{c|c} A_c & B_c \\ \hline -C_c & 0 \end{array}\right]$ is strictly positive real.

Note that since the plant is positive real and the negative feedback compensator is strictly positive real, condition (i) is automatically satisfied. Now, using the approach proposed in [34,38] we have the following result for constructing strictly positive real compensators. For convenience, define $V_1 \triangleq D_1 D_1^T$ and $V_2 \triangleq D_2 D_2^T$.

Theorem 6.3.[34,38]. Assume $G(s) \overset{\min}{\sim} \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & 0 \end{array}\right]$ is positive real, and let Q_0 and L satisfy (2.23) and (2.24) where Q_0 is positive definite. Furthermore, assume that there exist $2n \times 2n$ nonnegative-definite matrices Q and P satisfying

$$0 = \hat{A}Q + Q\hat{A}^T + V_1 - Q\hat{C}^T V_2^{-1} \hat{C} Q, \tag{6.24}$$

$$0 = \hat{A}^T P + P\hat{A} + R_1 - P\hat{B} R_2^{-1} \hat{B}^T P, \tag{6.25}$$

where R_1, R_2, V_1, and V_2 satisfy

$$V_1 = LL^T + \hat{B} R_2^{-1} \hat{B}^T > 0, \tag{6.26}$$

$$V_2 = R_2, \tag{6.27}$$

$$R_1 > \hat{C}^T R_2^{-1} \hat{C}. \tag{6.28}$$

Then the negative feedback compensator

$$-K(s) \sim \left[\begin{array}{c|c} A_c & B_c \\ \hline -C_c & 0 \end{array}\right] = \left[\begin{array}{c|c} \hat{A} - Q\hat{C}^T V_2^{-1} \hat{C} - \hat{B} R_2^{-1} \hat{B}^T P & Q\hat{C}^T V_2^{-1} \\ \hline R_2^{-1} \hat{B}^T P & 0 \end{array}\right] \tag{6.29}$$

is strictly positive real and satisfies the design criteria (i), (ii). Furthermore, the H_2 performance is given by

$$J(A_c, B_c, C_c) = \mathrm{tr}[QR_1 + Q\hat{C}^T V_2^{-1} \hat{C} QP]. \tag{6.30}$$

In order to compare the positive real majorant bounds developed in this and the previous section to the complex block-structured majorant bound [7] and the complex

structured singular value bound [41], next we provide a synopsis of the results developed in [7] and [41].

7. The Complex Block-Structured Majorant Bound and the Complex Structured Singular Value Bound

In this section we present a brief summary of the results from [7] involving the complex block-structured majorant bound and [1,41] involving the complex structured singular value bound. Consider the standard problem of a linear time-invariant dynamic system given by

$$z(s) = G_{11}(s)w(s) + G_{12}(s)u(s), \qquad (7.1)$$

$$y(s) = G_{21}(s)w(s) + G_{22}(s)u(s), \qquad (7.2)$$

$$u(s) = \Delta_s(s)y(s), \qquad (7.3)$$

where

$$\Delta_s(s) \in \Delta_s \triangleq \{\Delta_s : \Delta_s(\jmath\omega) \leq\leq \hat{\Delta}_s(\jmath\omega)\}. \qquad (7.4)$$

Note that (7.1)-(7.3) can be written as

$$\tilde{z}(s) = G(s)\tilde{w}(s), \qquad (7.5)$$

$$\tilde{w}(s) = \tilde{\Delta}(s)\tilde{z}(s) + \tilde{v}(s), \qquad (7.6)$$

where

$$G(s) \triangleq \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix}, \qquad \tilde{\Delta}(s) \triangleq \begin{bmatrix} 0 & 0 \\ 0 & \Delta_s(s) \end{bmatrix},$$

$$\tilde{z}(s) \triangleq \begin{bmatrix} z(s) \\ y(s) \end{bmatrix}, \quad \tilde{w}(s) \triangleq \begin{bmatrix} w(s) \\ u(s) \end{bmatrix}, \quad \tilde{v}(s) \triangleq \begin{bmatrix} w(s) \\ 0 \end{bmatrix}.$$

The system given by (7.5) and (7.6) could, for example, represent an uncertain system in a closed-loop configuration with the plant uncertainty Δ_s "pulled out" into a fictitious feedback loop as shown in Figure 1.

Next, we present a theorem which provides the complex block-structured majorant bound along with a robust stability condition for the linear time-invariant system (7.5) and (7.6).

Theorem 7.1.[7]. The feedback system given by (7.5) and (7.6) is robustly stable for all $\Delta_s \in \mathbf{\Delta}_s$ if

$$\rho[|G(\jmath\omega)|_M \hat{\Delta}(\jmath\omega)] < 1, \qquad \omega \in (0, \infty). \tag{7.7}$$

Furthermore, the output $\tilde{z}(\jmath\omega)$ satisfies the bound

$$|\tilde{z}(\jmath\omega)|_M \leq\leq [I - |G(\jmath\omega)|_M \hat{\Delta}(\jmath\omega)]^{-1} |G(\jmath\omega)\tilde{v}(\jmath\omega)|_M, \tag{7.8}$$

where

$$\hat{\Delta}(\jmath\omega) \triangleq \begin{bmatrix} 0 & 0 \\ 0 & \hat{\Delta}_s(\jmath\omega) \end{bmatrix}.$$

Next, we summarize the method for obtaining the complex structured singular value bound [41] for the standard problem addressed by (7.1)-(7.3). First recall that for complex multiple block-structured uncertainty $\Delta_s \in \mathbf{\Delta}_{bs}$, where

$$\mathbf{\Delta}_{bs} \triangleq \{\Delta_s : \Delta_s = \text{block--diag}(\Delta_1, \Delta_2, \cdots, \Delta_r), \ \Delta_i \in \mathbb{C}^{m_i \times m_i}, i = 1, \cdots, r\}, \tag{7.9}$$

and where $m_1, \cdots, m_r$ are given, the complex structured singular value $\mu_{\mathbf{\Delta}_{bs}}(G(\jmath\omega))$ is defined by

$$\mu_{\mathbf{\Delta}_{bs}}(G(\jmath\omega)) \triangleq \left(\min_{\Delta_s \in \mathbf{\Delta}_{bs}} \{\sigma_{\max}(\Delta_s) : \det(I - G(\jmath\omega)\Delta_s) = 0\} \right)^{-1}, \tag{7.10}$$

while $\mu_{\mathbf{\Delta}_{bs}}(G(\jmath\omega)) = 0$ if there exists no $\Delta_s \in \mathbf{\Delta}_{bs}$ such that $\det(I - G(\jmath\omega)\Delta_s) = 0$. It has been shown [1,41] that $\mu_{\mathbf{\Delta}_{bs}}(G(\jmath\omega))$ satisfies the inequality

$$\mu_{\mathbf{\Delta}_{bs}}(G(\jmath\omega)) \leq \inf_{N(\jmath\omega) \in \mathcal{N}_{\mathbf{\Delta}_{bs}}} \sigma_{\max}(N(\jmath\omega)G(\jmath\omega)N^{-1}(\jmath\omega)), \tag{7.11}$$

where $\mathcal{N}_{\mathbf{\Delta}_{bs}}$ denotes the set of positive-definite scaling matrices which are compatible with the uncertainty structure $\mathbf{\Delta}_{bs}$. Recall that if the number of blocks in Δ_s is three or less then the inequality in (7.11) is a strict equality [41].

As is well known [1,2,41], in order to consider robust performance within the complex structured singular value framework involving the standard problem given by (7.1)-(7.3), we introduce an additional uncertainty block Δ_p between w and z so that

$$w(s) = \Delta_p(s)z(s) \tag{7.12}$$

24

and require stability robustness in the face of all perturbations, including the fictitious block Δ_p. In this case, (7.1)-(7.3) along with (7.12) can be written as

$$\tilde{z}(s) = G(s)\tilde{w}(s), \tag{7.13}$$

$$\tilde{w}(s) = \Delta(s)\tilde{z}(s), \tag{7.14}$$

where

$$\Delta(s) \triangleq \begin{bmatrix} \Delta_p(s) & 0 \\ 0 & \Delta_s(s) \end{bmatrix}. \tag{7.15}$$

This feedback configuration can be captured by Figure 1 by setting $\tilde{v}$ to zero and replacing $\tilde{\Delta}(s)$ by $\Delta(s)$. Finally, note that the output $z(s)$ is related to the input $w(s)$ by

$$z(s) = \mathcal{G}(s)w(s), \tag{7.16}$$

where

$$\mathcal{G}(s) \triangleq \left[G_{11}(s) + G_{12}(s)\Delta_s(s)[I - G_{22}(s)\Delta_s(s)]^{-1}G_{21}(s)\right].$$

For the statement of the next result define

$$\boldsymbol{\Delta} \triangleq \{\Delta : \Delta = \text{block–diag}(\Delta_p, \Delta_s), \text{ and } \Delta_s \in \boldsymbol{\Delta}_{bs}\},$$

$$\mathbf{B}_{\boldsymbol{\Delta}_{bs}} \triangleq \{\Delta_s : \|\Delta_s\|_\infty \leq \frac{1}{\gamma} \text{ and } \Delta_s \in \boldsymbol{\Delta}_{bs}\},$$

$$\mathbf{B}_{\boldsymbol{\Delta}} \triangleq \{\Delta \in \boldsymbol{\Delta} : \|\Delta\|_\infty \leq \frac{1}{\gamma}\}.$$

Theorem 7.2. Let $\Delta \in \mathbf{B}_{\boldsymbol{\Delta}}$. Then the feedback system given by (7.13) and (7.14) is robustly stable for all $\Delta_s \in \boldsymbol{\Delta}_{bs}$ if and only if

$$\mu_{\boldsymbol{\Delta}_{bs}}(G_{22}(\jmath\omega)) < \gamma, \qquad \omega \in (0, \infty). \tag{7.17}$$

Furthermore, if $\mu_{\boldsymbol{\Delta}_{bs}}(G_{22}(\jmath\omega)) \leq \mu_{\boldsymbol{\Delta}}(G(\jmath\omega)) < \gamma$ then

$$\|z(\jmath\omega)\|_2 \leq \mu_{\boldsymbol{\Delta}}(G(\jmath\omega))\|w(\jmath\omega)\|_2. \tag{7.18}$$

Proof. The first part of the theorem involving robust stability is a direct consequence of Theorem 4.2 of [41]. Now, to show the performance bound (7.18) let $\mu_{\boldsymbol{\Delta}}(G(\jmath\omega)) = \beta < \gamma$,

25

so that $\det(I - G(\jmath\omega)\Delta) \neq 0$ for all $\|\Delta\|_\infty < \frac{1}{\beta}$. Furthermore if $\mu_{\Delta_{bs}}(G_{22}(\jmath\omega)) \leq \beta$, then $\det(I - G_{22}(\jmath\omega)\Delta_s) \neq 0$ for all $\|\Delta\|_\infty < \frac{1}{\beta}$. Now, since $\det(I - G(\jmath\omega)\Delta) = \det(I - G_{22}(\jmath\omega)\Delta_s)\det(I - \mathcal{G}(\jmath\omega)\Delta_p)$, it follows that $\det(I - \mathcal{G}(\jmath\omega)\Delta_p) \neq 0$, for all $\|\Delta\|_\infty < \frac{1}{\beta}$. Hence,

$$\max_{\Delta_s \in \Delta_{bs}, \|\Delta\|_\infty < \frac{1}{\beta}} \mu_{\Delta_p}(\mathcal{G}(\jmath\omega)) \leq \mu_\Delta(G(\jmath\omega)).$$

Furthermore, note that

$$\max_{\Delta_s \in B_{\Delta_{bs}}} \mu_{\Delta_p}(\mathcal{G}(\jmath\omega)) \leq \max_{\Delta_s \in \Delta_{bs}, \|\Delta\|_\infty < \frac{1}{\beta}} \mu_{\Delta_p}(\mathcal{G}(\jmath\omega)).$$

Now, using (7.16) the performance bound (7.18) is immediate. $\square$

Remark 7.1. Note that if, alternatively, $\mu_\Delta(G(\jmath\omega)) \leq \mu_{\Delta_{bs}}(G_{22}(\jmath\omega)) < \gamma$ is satisfied in the statement of Theorem 7.2 then (7.18) can be replaced by

$$\|z(\jmath\omega)\|_2 \leq \mu_{\Delta_{bs}}(G_{22}(\jmath\omega))\|w(\jmath\omega)\|_2.$$

Since our uncertainty characterization considered in Section 3 is in the time domain, we briefly outline a systematic approach to converting this uncertainty into the transfer function presented by the standard problem representation in (7.1)-(7.3). First, recall that an uncertain state-space model with disturbance $Dw(t)$ and performance variables $Ex(t)$ can be viewed as a system in a feedback configuration with the gain Δ_s (see Figure 1), that is,

$$\dot{x}(t) = Ax(t) + B_0 u(t) + Dw(t), \tag{7.19}$$

$$y(t) = C_0 x(t), \tag{7.20}$$

with feedback

$$u(t) = \Delta_s y(t), \tag{7.21}$$

and performance variables

$$z(t) = Ex(t). \tag{7.22}$$

26

Note that B_0 and C_0 in this formulation are fixed matrices denoting the structure of the uncertainty. Now, it follows that (7.19)-(7.22) yield

$$z(s) = E(sI - A)^{-1}Dw(s) + E(sI - A)^{-1}B_0u(s), \tag{7.23}$$

$$y(s) = C_0(sI - A)^{-1}Dw(s) + C_0(sI - A)^{-1}B_0u(s), \tag{7.24}$$

$$u(s) = \Delta_s(s)y(s), \tag{7.25}$$

which are equivalent in form to the standard problem equations given by (7.1)-(7.3). Now, with frequency and damping uncertainty, the system matrix in (3.1) can be written in second order canonical form $\hat{A} = \text{block-diag}(\hat{A}_i)$, $i = 1, \cdots, n$, where

$$\hat{A}_i = \begin{bmatrix} 0 & 1 \\ -\Omega_i^2 & -2\zeta_i\Omega_i \end{bmatrix},$$

or, equivalently,

$$\hat{A}_i = \begin{bmatrix} 0 & 1 \\ -\Omega_{0,i}^2 & -2\zeta_{0,i}\Omega_{0,i} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \gamma_i & \epsilon_i \end{bmatrix}, \quad i = 1, \cdots, n,$$

where $\Omega_{0,i}$ and $\zeta_{0,i}$ are the nominal natural frequencies and damping coefficients respectively, γ_i is the uncertainty in Ω_i^2, and ϵ_i is the uncertainty in $2\zeta_i\Omega_i$. Now, in order to use the above framework it need only be noted that

$$\begin{bmatrix} 0 & 0 \\ \gamma_i & \epsilon_i \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \gamma_i & 0 \\ 0 & \epsilon_i \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad i = 1, \cdots, n.$$

Of course, the above analysis also holds for a nominal closed-loop system with feedback uncertainty. In this case however, appropriate modifications to the system matrices in (7.23) and (7.24) are needed to capture the nominal closed-loop dynamics.

8. Illustrative Numerical Examples

In this section we present several illustrative numerical examples that demonstrate the effectiveness of the proposed positive real majorant bounds (PRMB) over the complex block-structured majorant bound (CBSMB) and the complex structured singular value bound (CSSVB).

Example 8.1. ($n = 1$, damping uncertainty)

27

Our first example considers performance bounding for the case $n = 1$ with one control input and damping uncertainty. The closed-loop system is given by (3.26) with

$$\Phi^{-1}(H_1, s) = \frac{s^2 + 2\zeta\Omega s + \Omega^2}{s},$$

$$F(H_2, s) = k_1 b^2,$$

$$k_1 = 2, \quad D = b = 1,$$

and with

$$\Omega = 10(2\pi)\frac{\text{rad}}{\text{sec}}, \quad \zeta = 0.01, \quad \widehat{\Delta\zeta} = 0.009.$$

In this case, Theorems 4.1 and 4.2 give the same performance bound which is shown in Figure 2. For this example, the complex block-structured majorant bound is totally non-conservative while the complex structured singular value bound gives the most conservative performance predictions.

Example 8.2. ($n = 1$, frequency uncertainty)

This example considers the same case as Example 8.1 except that the damping ratio is constant while the frequency is uncertain with

$$\widehat{\Delta\Omega} = 5(2\pi)\frac{\text{rad}}{\text{sec}}.$$

For the assumed uncertainty range both the complex block-structured majorant bound and the complex singular value bound are infinite since, in this case, both methods predict instability. The proposed positive real majorant bound gives a tight finite performance bound. This is shown in Figure 3.

Example 8.3. ($n = 3$, frequency uncertainty in the first mode)

This example considers performance bounds for three closely spaced modes with one control input and frequency uncertainty in the first mode. The closed-loop system is given by (3.26) with

$$\Phi^{-1}(H_1, s) = \text{diag}\{\frac{s^2 + 2\zeta_i\Omega_i s + \Omega_i^2}{s}\}_{i=1}^3,$$

$$F(H_2, s) = k_1 BB^{\mathrm{T}},$$

$$D = B,$$

$$k_1 = 2, \quad B = [1, 1, 1]^T,$$

28

and with

$$\{\Omega_1, \Omega_2, \Omega_3\} = \{10(2\pi)\frac{\text{rad}}{\text{sec}}, 15(2\pi)\frac{\text{rad}}{\text{sec}}, 50(2\pi)\frac{\text{rad}}{\text{sec}}\},$$

$$\{\zeta_1, \zeta_2, \zeta_3\} = \{0.005, 0.01, 0.0025\},$$

$$\{\widehat{\Delta\Omega_1}, \widehat{\Delta\Omega_2}, \widehat{\Delta\Omega_3}\} = \{5(2\pi)\frac{\text{rad}}{\text{sec}}, 0, 0\}.$$

Once again both the complex block-structured majorant bound and the complex structured singular value bound give infinite performance predictions. The positive real majorant bound shown in Figure 4 gives a finite performance bound. This bound was obtained by computing the minimum of the performance bounds given by Theorems 5.1 and 5.2 for each frequency.

Example 8.4. (n=3, frequency and mode shape uncertainty)

In order to compare the positive real majorant bounds obtained by Theorems 5.1 and 5.2 and Theorems 6.1 and 6.2 this example considers the same case as Example 8.3 with two sensors and actuators with both frequency and mode shape uncertainty. Specifically, the frequency uncertainty is as in Example 8.3 and the mode shape uncertainty with the assumed static controller are

$$B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad \widehat{\Delta B} = \begin{bmatrix} 0.01 & 0.01 \\ 0.01 & 0.01 \\ 0.01 & 0.01 \end{bmatrix}, \quad K = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}.$$

The corresponding bounds are shown in Figure 5. Note that since the assumed frequency range for this example is the same as the previous example the bounds of Section 7 do not give finite predictions.

Next, using the frequency domain performance bounds given by Theorems 6.1 and 6.2 along with Theorem 6.3 for constructing strictly positive real dynamic compensators we apply our results to a simply supported Euler-Bernoulli beam with multiple frequency uncertainty.

Example 8.5. Consider the simply supported Euler-Bernoulli beam with governing partial differential equation for the transverse deflection $w(x,t)$ given by

$$m(x)\frac{\partial^2 w(x,t)}{\partial t^2} + \frac{\partial^2}{\partial x^2}[EI(x)\frac{\partial^2 w(x,t)}{\partial x^2}] = f(x,t), \qquad (8.1)$$

29

and with boundary conditions

$$w(x,t)|_{x=0,L} = 0, \quad EI\frac{\partial^2 w(x,t)}{\partial x^2}\Big|_{x=0,L} = 0, \tag{8.2}$$

where $m(x)$ is mass per unit length and $EI(x)$ is the flexural rigidity with E denoting Young's modulus of elasticity and $I(x)$ denoting the cross-sectional area moment of inertia about an axis normal to the plane of vibration and passing through the center of the cross-sectional area. Finally, $f(x,t)$ is the force distribution due to control actuation and external disturbances. Assuming uniform beam properties, the modal decomposition of this system has the form

$$w(x,t) = \sum_{r=1}^{\infty} W_r(x)q_r(t), \tag{8.3}$$

$$\int_0^L mW_r^2(x)dx = 1, \quad W_r(x) = \sqrt{\frac{2}{mL}}\sin\frac{r\pi x}{L}, \quad r = 1, 2, \ldots, \tag{8.4}$$

where, assuming uniform proportional damping, the modal coordinates q_r satisfy

$$\ddot{q}_r(t) + 2\zeta\Omega_r\dot{q}_r(t) + \Omega_r^2 q_r(t) = \int_0^L f(x,t)W_r(x)dx, \quad r = 1, 2, \ldots. \tag{8.5}$$

For simplicity assume $L = \pi$ and $m = EI = 2/\pi$ so that $\sqrt{\frac{2}{mL}} = 1$. Furthermore, we place a colocated velocity/force actuator pair at $x = 0.55L$. Finally, modeling the first five modes and defining the plant state as $x = [q_1, \dot{q}_1, \cdots, q_5, \dot{q}_5]^T$, and defining the performance of the beam in terms of the velocity at $x = 0.7L$, the resulting state space model and problem data are

$$\hat{A} = \text{block-diag}_{i=1,\ldots,5}\begin{bmatrix} 0 & 1 \\ -\Omega_i^2 & -2\zeta\Omega_i \end{bmatrix}, \quad \Omega_i = i^2, \quad \zeta = 0.01,$$

$$\hat{B} = \hat{C}^T = [0 \quad 0.9877 \quad 0 \quad -0.309 \quad 0 \quad -0.891 \quad 0 \quad 0.5878 \quad 0 \quad 0.7071]^T,$$

$$E_1 = \begin{bmatrix} 0 & 0.809 & 0 & -0.951 & 0 & 0.309 & 0 & 0.5878 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$E_2 = [0 \quad 1.9]^T, \quad R_1 = E_1^T E_1, \quad D_1 = [\hat{B} \quad 0_{10\times1}], \quad D_2 = [0 \quad 1.9],$$

$$V_2 = R_2 = D_2 D_2^T = E_2^T E_2 = 3.61.$$

Using Thoerem 6.3 we design a strictly positive real dynamic compensator $K(s)$. Next, we assume frequency uncertainty in both Ω_1 and Ω_2 with $\widehat{\Delta\Omega}_1 = 0.5$ and $\widehat{\Delta\Omega}_2 = 0.5$. The

corresponding positive real majorant performance bound is given in Figure 6. Once again this bound was obatained by computing the minimum of the performance bounds given by Theorems 6.1 and 6.2 for each frequency. The complex block-structured majorant bound and the complex structured singular value bound predict instability for the assumed uncertainty range and hence give infinite performance bounds. In order to compare the performance bounds using all three methods, the maximum uncertainty range in the natural frequencies for which the two alternative methods guarantee stability is found. Specifically, complex structured singular value analysis predicts stability for the range of $\widehat{\Delta\Omega}_1 < 0.034$ and $\widehat{\Delta\Omega}_2 < 0.134$, while complex block-structured majorant analysis predicts stability for the range of $\widehat{\Delta\Omega}_1 < 0.071$ and $\widehat{\Delta\Omega}_2 < 0.144$. The parameter space for the above predictions is shown in Figure 7. Note that the positive real result guarantees unconditional stability. Now, with $\widehat{\Delta\Omega}_1 = 0.034$ and $\widehat{\Delta\Omega}_2 = 0.134$ which coressponds to the largest uncertainty range for which all three robustness tests guarantee stability, the comparision of the performance bounds for all three approaches is shown in Figure 8. Note that the proposed positive real majorant bound gives the tightest robust performance bound while the complex structured singular value bound (μ-performance bound) is the most conservative.

9. Conclusion

This paper developed frequency domain performance bounds for closed-loop systems consisting of positive real plants and strictly positive real compensators. The results are developed by using certain properties of the logarithmic norm in conjunction with majorant analysis. Unlike previous results in robustness analysis, the performance bounds remain finite even when the uncertainty is made large. The examples compared the new bounds with a previous majorant bound and the corresponding bound from complex structured singular value analysis. In all cases the new bound was much less conservative than the alternative bounds. Future work will involve extending these results to reduce the conservatism in the analysis of closed-loop systems for which the plant and controller are positive real only over a particular frequency band.

31
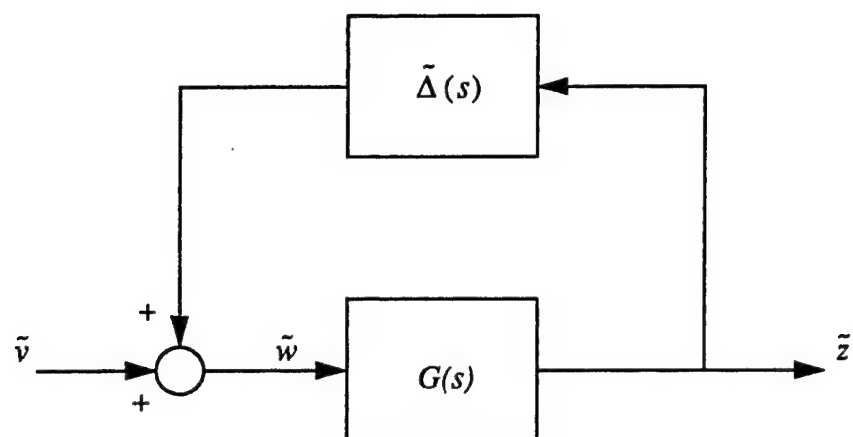
Figure 1. Nominal Closed-Loop with Feedback Uncertainty
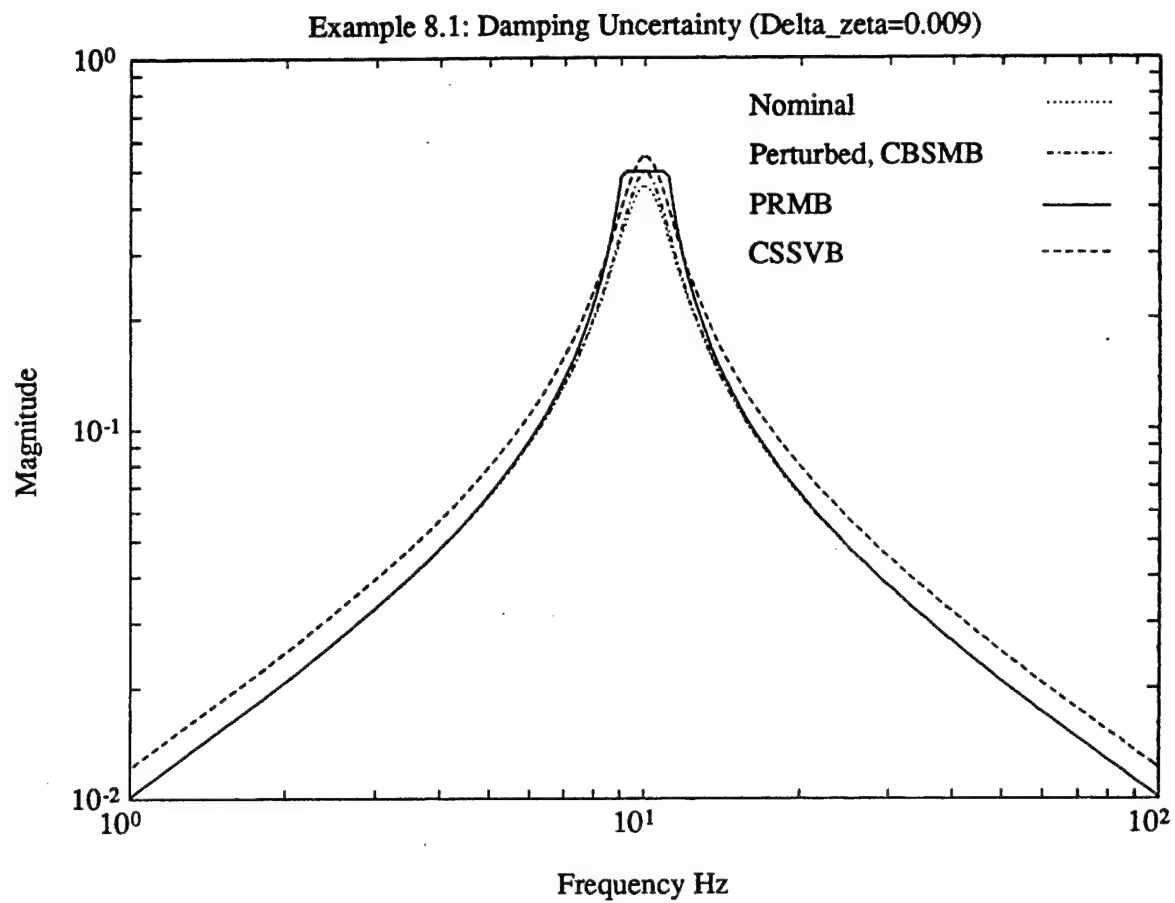
Figure 2. Performance Bounds for Example 8.1 (n=1, Damping Uncertainty)
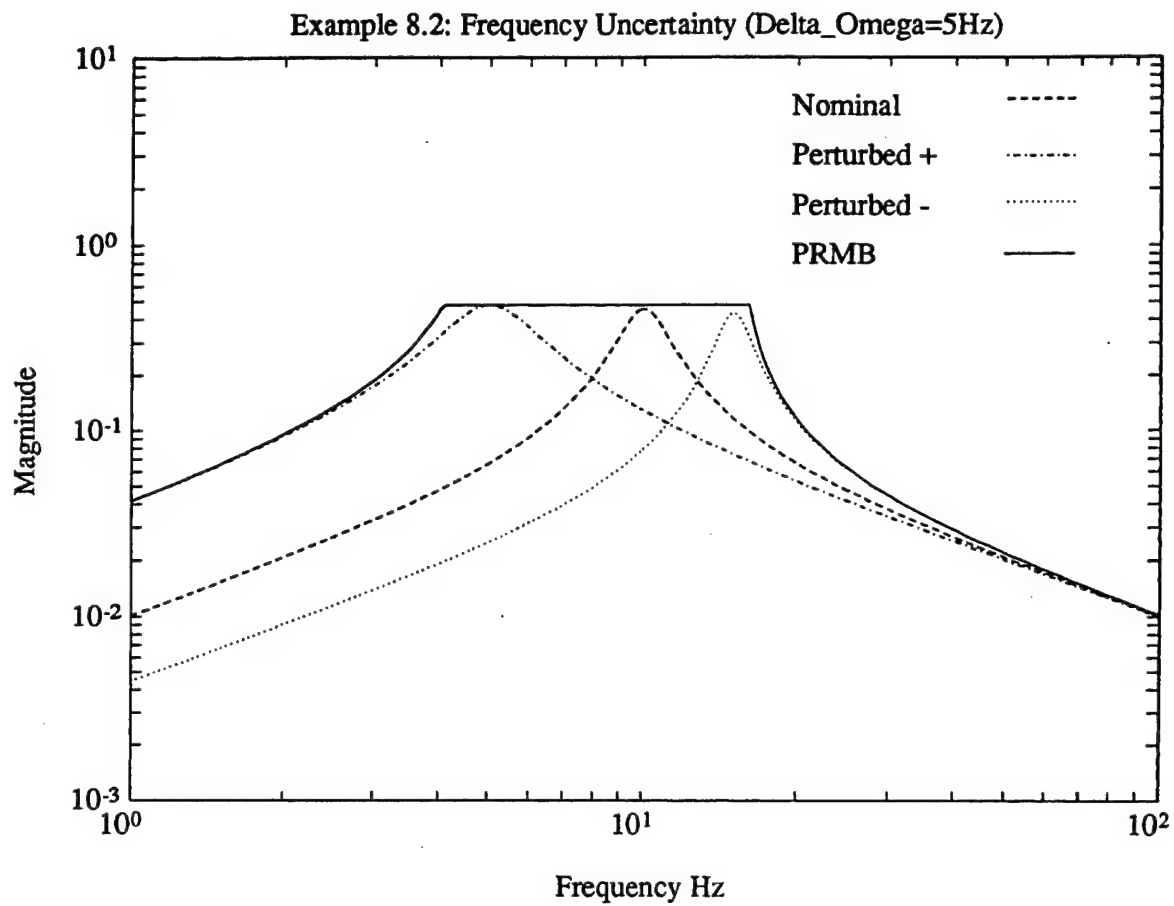
Figure 3. Performance Bound for Example 8.2 (n=1, Frequency Uncertainty)
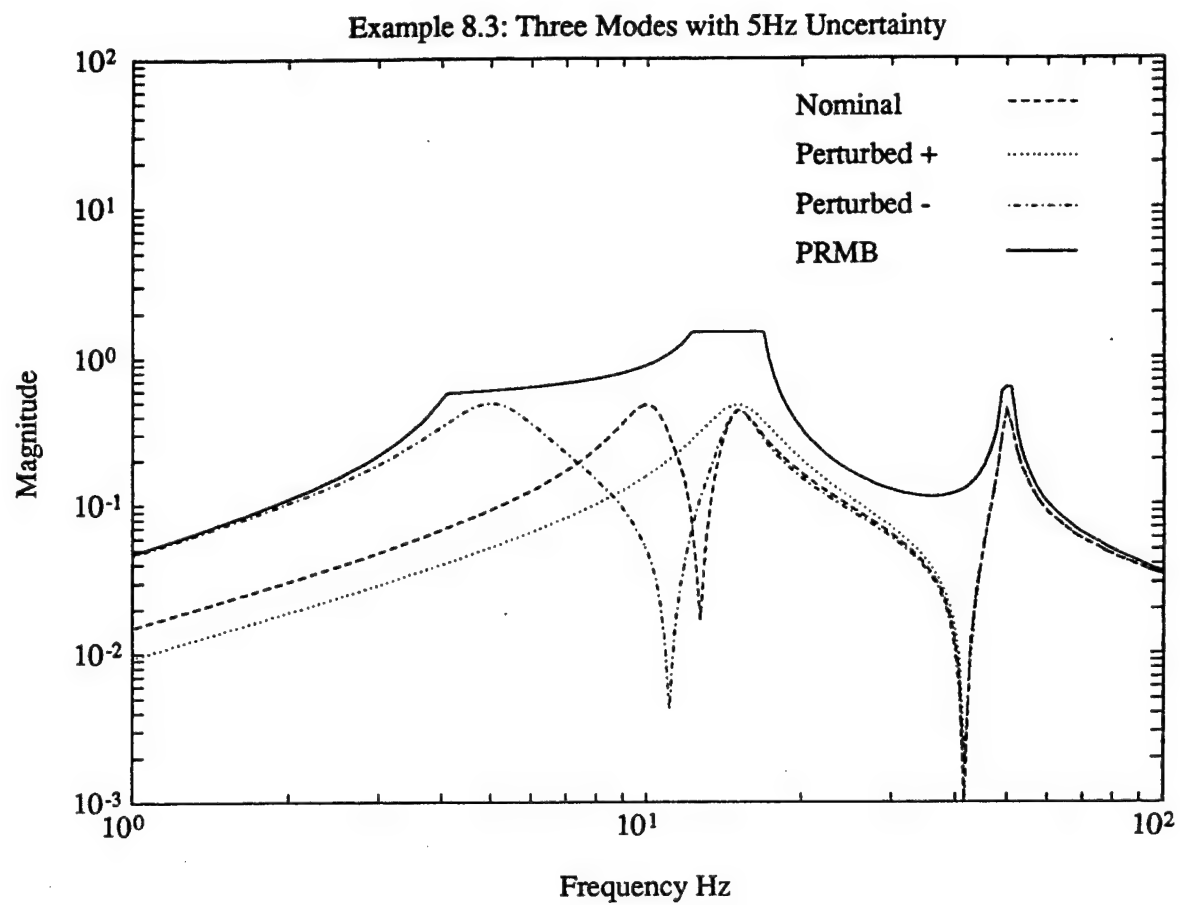
Figure 4. Performance Bound for Example 8.3 (n=3, Frequency Uncertainty)
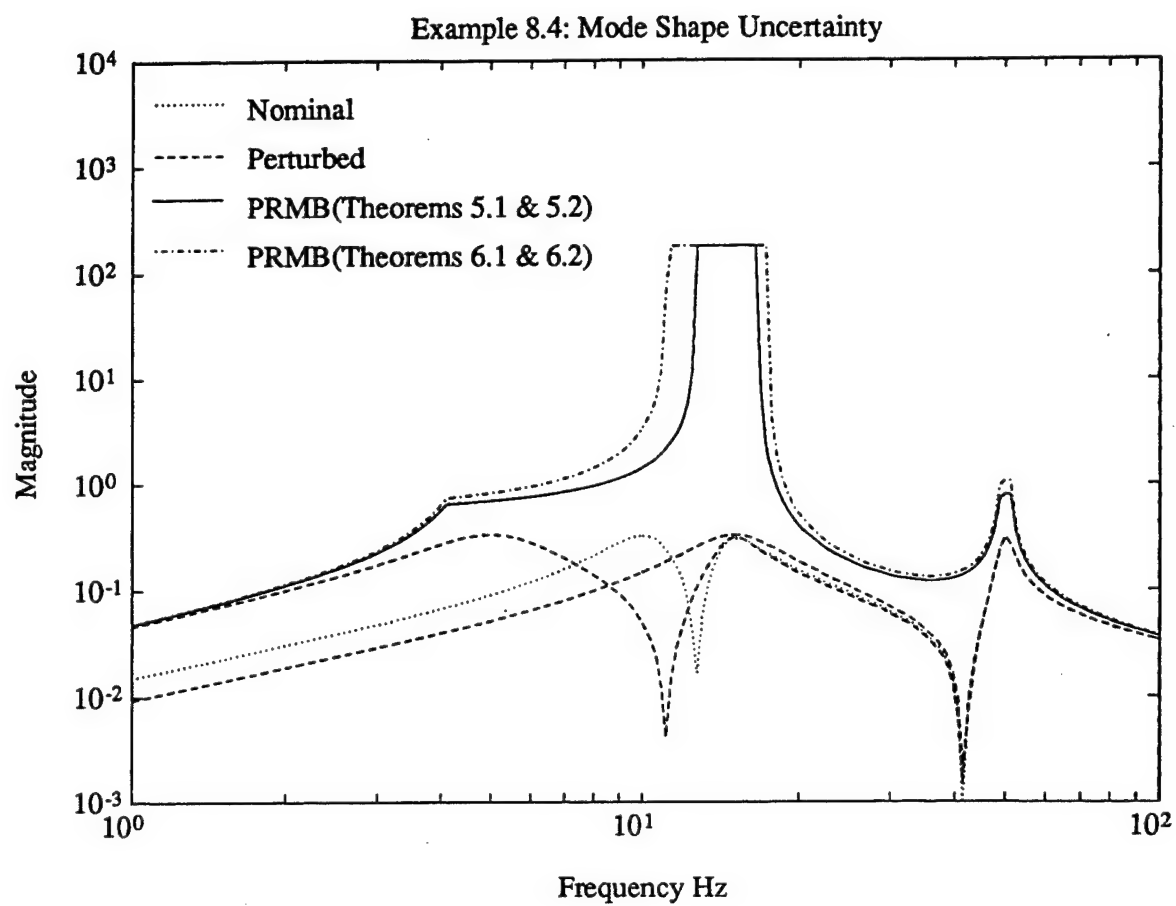
Figure 5. Comparision of Performance Bounds Obtained from
Theorems 5.1 and 5.2 and Theorems 6.1 and 6.2
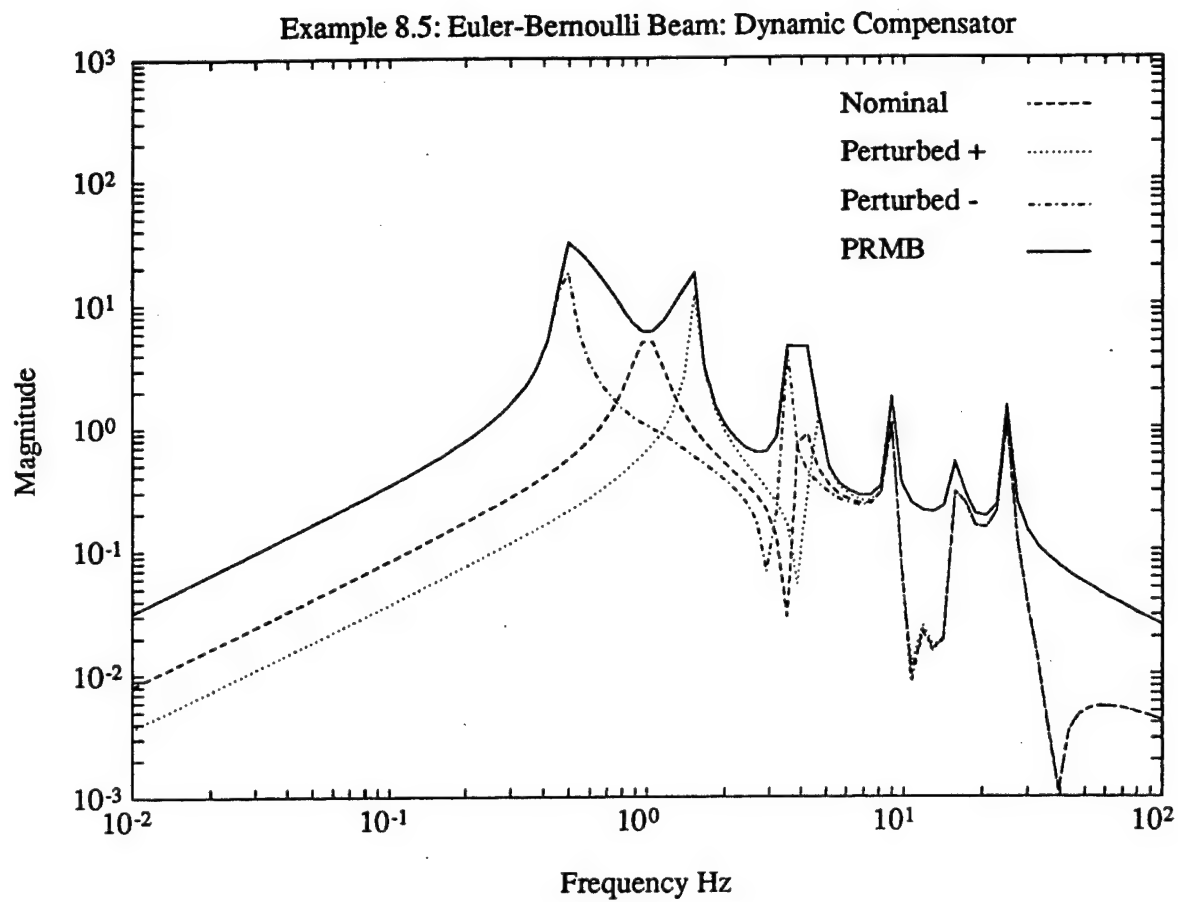
Figure 6. Performance Bound for the Euler-Bernoulli Beam (Example 8.5)
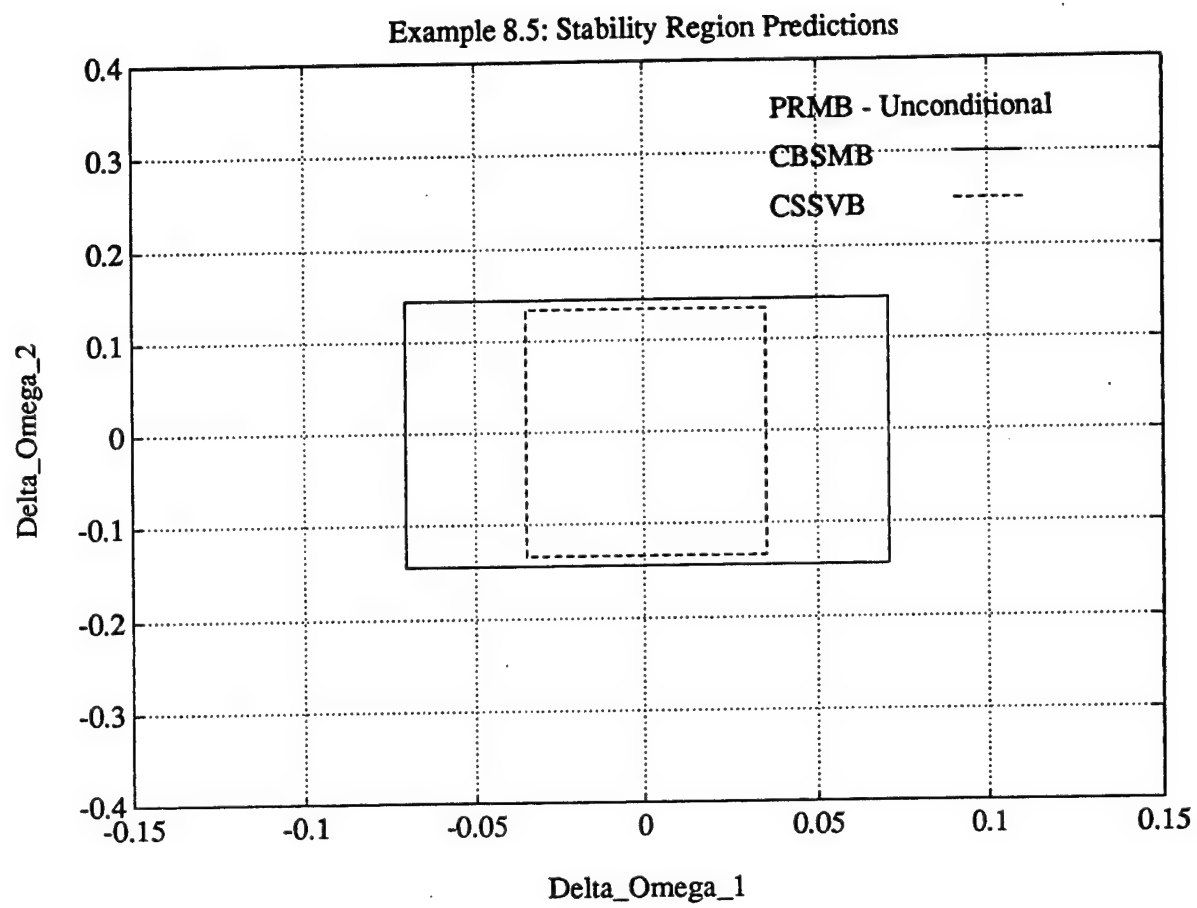
Figure 7. Guaranteed Stability Predictions for PRMB, CBSMB, CSSVB
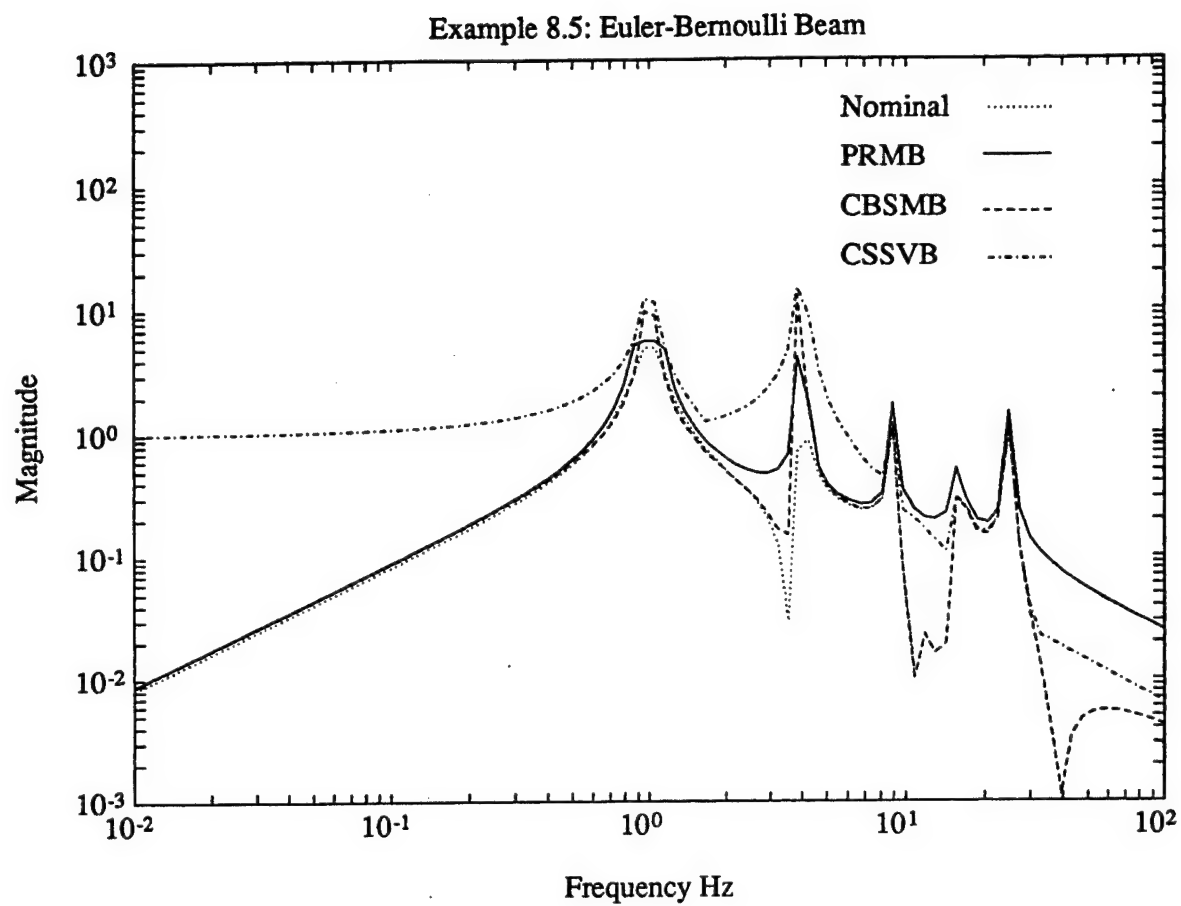
Figure 8. Comparision of PRMB, CBSMB, and CSSVB for the Euler-Bernoulli Beam Example

References

[1] J.C. Doyle, J.E. Wall, and G. Stein, "Performance and robustness analysis for structured uncertainty," in *Proc. IEEE Conf. Dec. Contr.*, Orlando, FL, pp. 629-636, 1982.

[2] J.C. Doyle, "Structured uncertainty in control system design," in *Proc. IEEE Conf. Dec. Contr.*, Ft. Lauderdale, FL, pp. 260-265, 1985.

[3] R.K. Yedavalli, "Time domain robust control design for linear quadratic regulators by perturbation bound analysis," in *Proc. IFAC Workshop on Model Error Concepts and Compensation*, R. E. Skelton and D. H. Owens, Eds., Boston, MA, pp. 129-135, 1985.

[4] R.D. Johnston and G.W. Barton, "Control system performance robustness," *Int. J. Contr.*, Vol. 45, pp. 641-648, 1987.

[5] D.C. Hyland and D.S. Bernstein, "The majorant Lyapunov equation: A nonnegative matrix equation for guaranteed robust stability and performance of large scale systems," *IEEE Trans. Autom. Contr.*, Vol. AC-32, pp. 1005- 1013, 1987.

[6] E.G. Collins, Jr. and D.C. Hyland, "Improved robust performance bounds in covariance majorant analysis," *Int. J. Contr.*, Vol. 50, pp. 495-509, 1989.

[7] D.C. Hyland and E.G. Collins, Jr., "An M-matrix and majorant approach to robust stability and performance analysis for systems with structured uncertainty," *IEEE Trans. Autom. Contr.*, Vol. AC-34, pp. 691-710, 1989.

[8] D.C. Hyland and E.G. Collins, Jr., "Some majorant robustness results for discrete-time systems," *Automatica*, Vol. 27, pp. 167-172, 1991.

[9] D.S. Bernstein and W.M. Haddad, "Robust stability and performance analysis for state space systems via quadratic Lyapunov bounds," *SIAM J. Matrix Anal. Appl.*, Vol. 11, pp. 239-271, 1990.

[10] D.S. Bernstein and W.M. Haddad, "Robust stability and performance analysis for linear dynamic systems," *IEEE Trans. Autom. Contr.*, Vol. AC-34, pp. 751-758, 1989.

[11] W.M. Haddad and D.S. Bernstein, "Parameter-dependent Lyapunov functions, constant real parameter uncertainty, and the Popov criterion in robust analysis and synthesis: Part 1," in *Proc. IEEE Conf. Dec. Contr.*, Brighton, U.K., pp. 2274-2279, 1991.

[12] W.M. Haddad and D.S. Bernstein, "Parameter-dependent Lyapunov functions, constant real parameter uncertainty, and the Popov criterion in robust analysis and synthesis: Part 2," in *Proc. IEEE Conf. Dec. Contr.*, Brighton, U.K., pp. 2618-2623, 1991.

[13] W.M. Haddad, J.P. How, S.R. Hall, and D.S. Bernstein, "Extensions of mixed-μ bounds to monotonic and odd monotonic nonlinearities using absolute stability theory," *Int. J. Contr.*, to appear.

[14] A.M. Ostrowski, "On some metrical properties of operator matrices and matrices partitioned into blocks," *J. Math. Anal. Appl.*, Vol. 10, pp. 161-209, 1975.

[15] D.C. Hyland and E.G. Collins, Jr., "Block Kronecker products and block norm matrices in large-scale systems analysis," *SIAM J. Matrix Anal. Appl.*, Vol. 10, pp. 18-29, 1989. Erratum, Vol. 10, p. 593, 1989.

[16] G. Dahlquist, "On matrix majorants and minorants with applications to differential equations," *Lin. Alg. Appl.*, Vol. 52/53, pp. 199-216, 1983.

[17] A.D.W. Porter and A.N. Michel, "Input-output stability of time-varying nonlinear multiloop feedback systems," *IEEE Trans. Autom. Contr.*, Vol. AC-19, pp. 422-427, 1974.

[18] B.E.L. Lasley and A.N. Michel, "Input-ouput stability of interconnected systems," *IEEE Autom. Trancs. Contr.*, Vol. AC-21, pp. 84-89, 1976.

[19] T. Strom, "On logarithmic norms," *SIAM J. Num. Anal.*, Vol. 12, pp. 741-753, 1975.

[20] M. Fiedler and V. Ptak, "On matrices with non-positive off-diagonal elements and positive principal minors," *Czechoslovakian Math. J.*, Vol. 12, pp. 382-400, 1962.

[21] E. Seneta, *Non-Negative Matrices*. New York: Wiley, 1973.

[22] A. Berman and R.J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. New York: Academic, 1979.

[23] R.A. Horn and R.C. Johnson, *Matrix Analysis*. New York: Cambridge University Press, 1985.

[24] B.D.O. Anderson and S. Vongpanitlerd, *Network Analysis and Synthesis: A Modern Systems Theory Approach*. Englewood Cliffs: Prentice-Hall, 1973.

[25] R.J. Benhabib, R.P. Iwens, and R.L. Jackson, "Stability of large space structure control systems using positivity concepts," *J. Guid. Contr.*, Vol. 4, pp. 487-494, 1981.

[26] R. Lozano-Leal and S. Joshi, "Strictly positive real transfer functions revisited," *IEEE Trans. Autom. Contr.*, Vol. AC-35, pp. 1243-1245, 1990.

[27] J.T. Wen, "Time domain and frequency domain conditions for strict positive realness," *IEEE Trans. Autom. Contr.*, Vol. AC-33, pp. 988-992, 1988.

[28] B.D.O. Anderson, "A system theory criterion for positive real matrices," *SIAM J. Contr. Optim.*, Vol. 5, pp. 171-182, 1967.

[29] B.D.O. Anderson, "Dual form of a positive real lemma," *Proc. IEEE*, Vol. 55, pp. 1749-1750, 1967.

[30] W.M. Haddad and D.S. Bernstein, "Robust stabilization with positive real uncertainty: beyond the small gain theorem," *Sys. Contr. Let.*, Vol. 17, pp. 191-208, 1991.

[31] M.J. Balas, "Direct velocity feedback control of large space structures," *J. Guid. Contr.*, Vol. 2, pp. 252-253, 1967.

[32] S.M. Joshi, "Robustness properties of collocated controllers for flexible spacecraft," *J. Guid. Contr.*, Vol 9, pp. 85-91, 1986.

[33] M.D. McLaren and G.L. Slaterr, "Robust multivariable control of large space structures using positivity," *J. Guid. Contr. Dyn.*, Vol. 10, pp. 393-400, 1987.

[34] R. Lozano-Leal and S.M. Joshi, "On the design of dissipative LQG-type controller," in *Proc. Conf. Dec. Contr.*, Austin, TX, pp. 1645-1646, 1988. Also in P. Dorato and R.K. Yedavalli, *Recent Advances in Robust Control.* New York: IEEE Press, 1990.

[35] G. Hewer and C. Kenney, "Dissipative LQG control systems," *IEEE Trans. Autom. Contr.*, Vol. AC-34, pp. 866-870, 1989.

[36] M.J. Jacobus, M. Jamshidi, C. Abdallah, P.Dorato, and D.S. Bernstein, "Design of strictly positive real, fixed-order dynamic compensators," in *Proc. IEEE Conf. Dec. Contr.*, Honolulu, HI, pp. 3492-3495, 1990.

[37] J.D. Gardiner, "Stabilizing control for second-order models and positive real systems," *AIAA J. Guid. Contr. Dyn.*, Vol. 15, pp. 280-282, 1992.

[38] W.M. Haddad, D.S. Bernstein, and Y.W. Wang, "Dissipative H_2/H_∞ controller synthesis," *IEEE Trans. Autom. Contr.*, to appear.

[39] M.G. Safonov, E.A. Jonckheere, and D.J.N. Limebeer, "Synthesis of positive real multivariable feedback systems," *Int. J. Contr.*, Vol. 45, pp. 817-842, 1987.

[40] D.G. MacMartin and S.R. Hall, "Control of uncertain structures using an H_∞ power flow approach," *J. Guid. Contr. Dyn.*, Vol. 14, pp. 521-530, 1991.

[41] A. Packard and J.C. Doyle, "The complex structured singular value," *Automatica*, Vol. 29, pp. 71-109, 1993.

Appendix G:

Optimal Popov Controller Analysis and Synthesis for Systems with Real Parameter Uncertainty

Optimal Popov Controller Analysis and Synthesis for Systems with Real Parameter Uncertainties*

by

Jonathan P. How
Space Engineering Research Center
Massachusetts Institute of Technology
Room 37-375, Cambridge, MA 02139
Tel: (617) 253-4923
Email: jon@sputnik.mit.edu

Emmanuel G. Collins, Jr.
Harris Corporation
Government Aerospace Systems Division
MS 19/4849
Melbourne, FL 32902
Email: ecollins@x102a.ess.harris.com

Wassim M. Haddad
Department of Mechanical and
Aerospace Engineering
Florida Institute of Technology
Melbourne, FL 32901
Tel: (407) 768-8000 Ext. 7241
Email: haddad@zach.fit.edu

Keywords: Robust performance analysis and synthesis,
              control system design, real parameter uncertainty.

Running Title: Popov Controller Analysis and Synthesis.

Abstract

Robust performance analysis is very important in the design of controllers for uncertain multivariable systems. Recent research has investigated the use of absolute stability criteria to develop less conservative analysis tests for systems with linear and nonlinear real parameter uncertainties. This paper extend previous work on optimal $\mathcal{H}_2$ performance analysis with the Popov criterion. In particular, an algorithm is presented that can be used to analyze systems with multiple uncertainties that have both upper and lower robustness bounds. More general Popov stability multipliers and less restrictive assumptions on the structure of the uncertainty block are also included. The analysis is performed using a numerical homotopy algorithm. The technique is demonstrated on robust compensators that have been designed for the Middeck Active Control Experiment (MACE): a Shuttle program scheduled for flight in December, 1994. The analysis clearly shows the relative robustness capabilities of the robust controllers used in the iterative control design methodology that has been developed for the uncertain dynamics of MACE. The analysis is also combined with Popov controller synthesis to yield a more sophisticated design technique for compensators that provide guaranteed robust performance.

2

The Popov analysis and synthesis algorithms have been developed from research on the Popov stability criterion [12] from *absolute stability theory* [11,13–17]. The absolute stability criteria provide sufficent conditions for the stability of a system in feedback with a particular class of sector-bounded, static nonlinear functions [12,17]. Note that a class of nonlinear functions can be associated with a set of system uncertainties considered in robust control. Thus, the Popov controller synthesis approach to robust control directly considers nonlinear real parameter uncertainties and treats linear uncertainties as a subset of this much broader class.

The state space tests from absolute stability theory with Lur'e-Postnikov Lyapunov functions are well documented [17], but it is only recently that the significance of the parameterized Lyapunov functions for robust control, in terms of a restriction on the time variation of the uncertainty set, has been understood [11,15,14,18]. A frequency domain representation of the absolute stability criteria is used in Refs. [19,14,18,17,20] to demonstrate that the robustness tests include magnitude and phase information about the system uncertainties. Both characteristics of the uncertainty must be considered to develop nonconservative tests for a system with real parameter uncertainty that is restricted to have phase of $\pm 180°$.

The primary purpose of this paper is to present several advances in Popov controller analysis and synthesis. Examples in Ref. [21] illustrate that the state space Popov analysis criterion is much less conservative than similar positive real and small gain ($\mathcal{H}_\infty$) criteria. This paper extends these previous results by considering systems with multiple uncertainties that have both upper and lower sector-bounds. The stability criterion is also developed using a more general stability multiplier

$$W(s) = H + Ns, \tag{1}$$

where $N \geq 0$, and $H > 0$ is not restricted to be the identity matrix, as required in Ref. [21]. Furthermore, the algorithm is developed with fewer restrictions on the structure of the system uncertainty.

The optimal Popov analysis algorithm is demonstrated using several robust controllers that were designed using a finite element model of MACE [2,22]. The results clearly show the relative robustness capabilities of the various techniques, and thus further illustrate their role in the iterative control design methodology discussed earlier. The best of these controllers was refined using the optimal Popov controller synthesis algorithm developed in Ref. [11,19,23]. Together, the two algorithms combine an improved analysis capability with a synthesis technique that guarantees robust performance for systems with real parameter uncertainty. In the process, this combination overcomes one of the main difficulties with the original synthesis algorithm: developing the stability multipliers for large guaranteed stability bounds [11,19].

3

DTIC COULD NOT GET MISSING

PAGES FROM CONTRIBUTOR

4

DTIC COULD NOT GET MISSING

PAGES FROM CONTRIBUTOR

5

DTIC COULD NOT GET MISSING

PAGES FROM CONTRIBUTOR

6

have both upper and lower sector-bounds. The structure of these uncertainties is also made less restrictive by removing the assumption that $C_0 B_0 = 0$. This section provides an outline of the homotopy algorithm, and an example of the approach is presented in the last section.

For clarity, the cost in Theorem 1 is rewritten here as

$$\overline{J} = \frac{1}{\alpha} \text{tr} \left[(P + C_0^T (M_2 - M_1) N C_0) V_{xx} \right], \tag{16}$$

where P is the solution of

$$0 = A_m^T P + P A_m + (\overline{C} - B_0^T P)^T R_0^{-1} (\overline{C} - B_0^T P) + \alpha R_{xx}, \tag{17}$$

and

$$A_m \triangleq A - B_0 M_1 C_0, \tag{18}$$

$$\overline{C} \triangleq H C_0 + N C_0 A_m, \tag{19}$$

$$C_{\text{ricc}} \triangleq \overline{C} - B_0^T P, \tag{20}$$

$$A_1 \triangleq A_m - B_0 R_0^{-1} \overline{C}, \tag{21}$$

$$A_2 \triangleq A_m - B_0 R_0^{-1} C_{\text{ricc}}. \tag{22}$$

Note that, with these definitions, Eq. 17 can be rewritten in the more familiar form

$$0 = A_1^T P + P A_1 + \overline{C}^T R_0^{-1} \overline{C} + \alpha R_{xx} + P B_0 R_0^{-1} B_0^T P. \tag{23}$$

The Lagrangian ($\mathcal{L}$) for the system is then formed by combining the cost overbound in Eq. 16 with the constraint in Eq. 17 using Lagrange multiplier matrix Q. The derivatives of this Lagrangian with respect to the free parameters in the design are the first-order necessary conditions that must be satisfied to determine an optimal solution. In particular, differentiating with respect to P yields a Lyapunov equation for the Lagrange multiplier matrix Q

$$0 = A_2 Q + Q A_2^T + \frac{1}{\alpha} V_{xx}. \tag{24}$$

Note that, if $(\cdot)$ refers to any free parameter in the optimization process, then $\partial \mathcal{L} / \partial (\cdot) = \partial \overline{J} / \partial (\cdot)$ [27]. The optimization problem then is to find values of α, H, and N that satisfy

$$\frac{\partial \overline{J}}{\partial \alpha} \triangleq H_1 = \text{tr} \left[Q R_{xx} \right] - \frac{1}{\alpha} \overline{J} = 0, \tag{25}$$

$$\frac{\partial \overline{J}}{\partial N} \triangleq H_2 = \frac{1}{\alpha} (M_2 - M_1) C_0 V_{xx} C_0^T + 2 R_0^{-1} C_{\text{ricc}} Q A_2^T C_0^T = 0, \tag{26}$$

$$\frac{\partial \overline{J}}{\partial H} \triangleq H_3 = 2 R_0^{-1} C_{\text{ricc}} Q (C_0 - (M_2 - M_1)^{-1} R_0^{-1} C_{\text{ricc}})^T = 0, \tag{27}$$

where P and Q are the solutions of Eq. 17 and Eq. 24, respectively. Note that only the elements of H_2 and H_3 corresponding to free parameters in N and H can be set to zero in the optimization process [11, 15, 18].

DTIC COULD NOT GET MISSING

PAGES FROM CONTRIBUTOR

8

Note that $M_2(0) = M_{2o}$ and $M_2(1) = M_{2f}$. The overall goal is to solve the optimization problem at $\lambda = 1$. However, it is usually quite difficult to develop valid initial conditions for the optimization at this point. Thus, a starting point that is simpler to initialize is introduced, and this corresponds to $\lambda = 0$ in Eq. 33. The purpose of the homotopy procedure is to obtain a solution to the optimization problem at $\lambda = 1$ by starting at $\lambda = 0$ and predicting the solution at $\lambda + \Delta\lambda$ based on the system derivatives and the solution at λ [21].

The continuation map is defined as the gradient of the upper bound on the cost function for the homotopy parameters M_1 and M_2. To compute this map, define

$$L(\eta, \lambda) \triangleq \begin{bmatrix} H_1(\eta, \lambda) \\ \text{vec}_N(H_2(\eta, \lambda)) \\ \text{vec}_N(H_3(\eta, \lambda)) \end{bmatrix}. \tag{34}$$

Then, as indicated by Eqs. 25–27, the continuation curve is given by $L(\eta, \lambda) = 0$ for $\lambda \in [0, 1]$. Then, taking both α and θ to be functions of λ, we can differentiate $L(\eta(\lambda), \lambda) = 0$ with respect to λ to yield *Davidenko's differential equation* [28, 29]

$$\frac{\partial L}{\partial \eta} \frac{d\eta}{d\lambda} + \frac{\partial L}{\partial \lambda} = 0. \tag{35}$$

Together with $\eta(0) = \eta_0$, this differential equation defines an initial value problem which, by numerical integration from $\lambda = 0$ to $\lambda = 1$, yields the desired solution $\eta(1)$ (see Ref. [26] for further details). As indicated by Eq. 35, the solution algorithm requires the computation of the Jacobian of $L(\alpha, \theta, \lambda)$, which is given by

$$\nabla L(\alpha, \theta, \lambda)^T = \begin{bmatrix} L_\eta & L_\lambda \end{bmatrix}, \tag{36}$$

where

$$L_\eta \triangleq \frac{\partial L}{\partial \eta}, \quad L_\lambda \triangleq \frac{\partial L}{\partial \lambda}. \tag{37}$$

The expressions for these gradients are given in the Appendix. The combined prediction–correction sequence is presented in Table 1. The algorithm starts with a correction step that updates the initial guess. When this step has converged, the value of λ is increased, and the changes in the multiplier coefficients are predicted in step 3d. The value of λ is increased until $\|L(\alpha, \theta, \lambda)\|$ is larger than a specified tolerance. The predictions of the coefficients η are then corrected in step 4. The cycle is repeated until $\lambda = 1$ or the stability prediction limit is reached. The output from the program is a list of optimal multipliers and the cost overbound at several values of λ, which define the different guaranteed stability regions.

The procedure in Table 1 is initialized by finding a set of scaling parameters for a given controller so that acceptable solutions exist for Eqs. 17 and 24. This step is typically one of the most difficult parts of a gradient search algorithm such as the one described in this paper.

9

DTIC COULD NOT GET MISSING

PAGES FROM CONTRIBUTOR

10
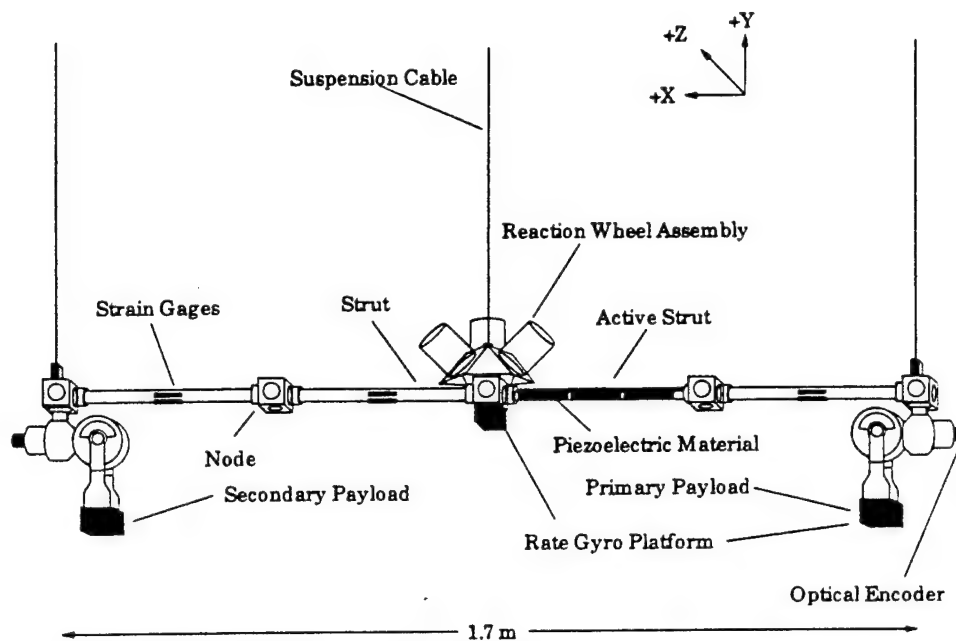
Figure 2: Middeck Active Control Experiment (MACE) EM test article suspended in 1-g by a three point suspension system.

ity active structural control in zero gravity conditions. The prediction of on-orbit closed-loop dynamics is based on analysis and ground testing. The goal of active control is to maintain pointing accuracy of one payload, while the remaining payload is moving independently. References [1, 2, 30] describe the experiment goals, hardware, and analytical modelling in some detail.

The configuration of the MACE test article was chosen to be representative of precision controlled, high payload mass fraction spacecraft, such as Earth observing platforms, with multiple, independently pointing or scanning payloads [31]. The Engineering Model (EM) is shown in Figure 2. Note that the X, Y, and Z axes are horizontal, vertical, and into the figure, respectively. The test article consists of a flexible bus to which are mounted two payloads, a reaction wheel assembly for attitude control, and various other sensors and actuators. Each payload is mounted to the structure using a two-axis gimbal that provides pointing capability. The EM is instrumented with angle encoders on each gimbal axis, a three axis rate gyro platform mounted under the reaction wheel assembly, and a two axis rate gyro platform mounted in the primary payload. The bus is composed of circular cross-section Lexan™ struts connected by aluminum nodes. The structure is supported for ground tests by a pneumatic/electric low-frequency suspension system [32].

As discussed in Ref. [2], several robust controllers were designed for the MACE test article using the finite element model discussed in Refs. [22, 33]. Control experiments have been performed for the full XYZ dynamics of MACE. However, for this paper, we consider

11

DTIC COULD NOT GET MISSING

PAGES FROM CONTRIBUTOR

12

For this closed-loop system, the cost function is determined by the nonnegative definite matrices $R_{xx} \triangleq C_{cl}^T C_{cl}$ and $V_{xx} \triangleq B_{cl} B_{cl}^T$. If the uncertainty in the open-loop system is represented by $A_{ol} + \Delta A_{ol}$, then the model error for the closed-loop system is written as

$$\Delta A_{cl} \triangleq \begin{bmatrix} \Delta A_{ol} & 0_{n \times n_c} \\ 0_{n_c \times n} & 0_{n_c \times n_c} \end{bmatrix}. \tag{45}$$

The nominal system dynamics and model error in Eq. 42–45 can then be combined to write the actual closed-loop system dynamics in the form of Eqs. 2 and 3.

The advantages and disadvantages of using finite element models (FEM) for MACE control design are discussed in Refs. [2, 3, 33]. The primary advantage is that this analytical modelling method can be used to predict the on-orbit dynamics prior to launch. However, a key difficulty is that FEM's tend to be much less accurate than equivalently sized measurement models [34]. In particular, there tend to be substantial errors in some modal frequencies and damping ratios, which is certainly true for a complicated structure such as MACE. Note that, to retain the capability of predicting the 0-g dynamics of the MACE test article, the measured 1-g data is used to update the physical parameters of the test article model, and not just the frequency characteristics of a particular state space model. This update procedure is a very difficult task, and is the subject of ongoing research [33]. Thus, the work in Ref. [2] used the FEM as given, and no attempt was made to update the state space model to account for obvious damping or frequency errors, because this would inconsistent with the purpose of the FEM in the MACE project [1].

For a system with many model errors, it is important to determine those that are most critical to the control design. As discussed in Refs. [2,3], the most important uncertainties to consider can be determined using a combination of analysis techniques based on the singular values of the sensitivity function, the multivariable Nichols test, and preliminary experiments. These techniques were used to determine that the frequencies of three modes were the most important uncertainties for these control designs. Note that these errors correspond to real parameter uncertainties in the system model.

The errors in these three modes are listed in Table 2. The finite element model results are compared with an identified model from measured data [34]. The "violin" description refers to modes with substantial interaction between the test article and the suspension system. Note that experience from these and other experiments has demonstrated that small errors in the damping values tend to degrade closed-loop performance, but small errors in the natural frequencies often lead to closed-loop instability. Thus, the model errors considered here address the more important issue of modal frequency uncertainty. For future comparison, note that the lowest frequency mode of these three was nearly destabilized by a very low authority LQG design that achieved only a 4 dB performance improvement.

13

14
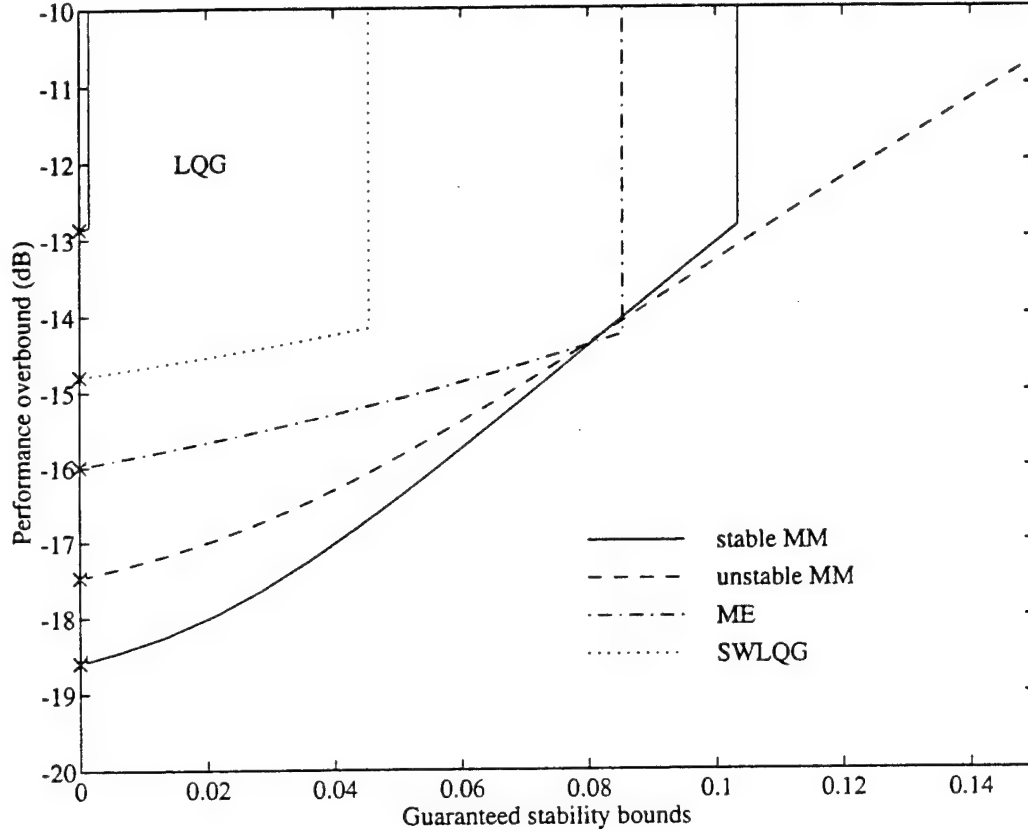
Figure 3: Robust stability and performance analysis using several controllers for MACE. Symbols × indicate nominal $\mathcal{H}_2$ performance for each design.

perturbation factor. The actual dynamics for each uncertain mode can be written as $\hat{A}_i = A_i + \triangle A_i$, where

$$\triangle A_i = \delta_i B_{0i} C_{0i}, \quad B_{0i} = \begin{bmatrix} 0 \\ -\omega_i \end{bmatrix}, \text{ and } C_{0i} = [\, \omega_i \quad 0\,], \quad i = 1, \ldots, 3. \tag{47}$$

The B_0 and C_0 matrices for the open-loop system are constructed from these two sets and then represent the structure of the uncertainty in an internal feedback model. These matrices are augmented with additional zeroes to compensate for the dynamics of the controller [19, 15], and they then can be used to represent the structure of the uncertainty in the closed-loop system $\triangle A_{cl}$.

Note that $C_0 B_0 = 0$ with the uncertainty model in Eq. 47. However, the code used to perform this example was written for the more general structure of $\triangle A$ with $C_0 B_0 \neq 0$. Note that removing the assumption in Ref. [21] significantly complicated the expressions for P in Eq. 14 and R_0 in Eq. refeq:4.4. As indicated in Section 4, the code could also be used for the block diagonal multipliers that are associated with repeated parameter uncertainties. This

15

DTIC COULD NOT GET MISSING

PAGES FROM CONTRIBUTOR

16
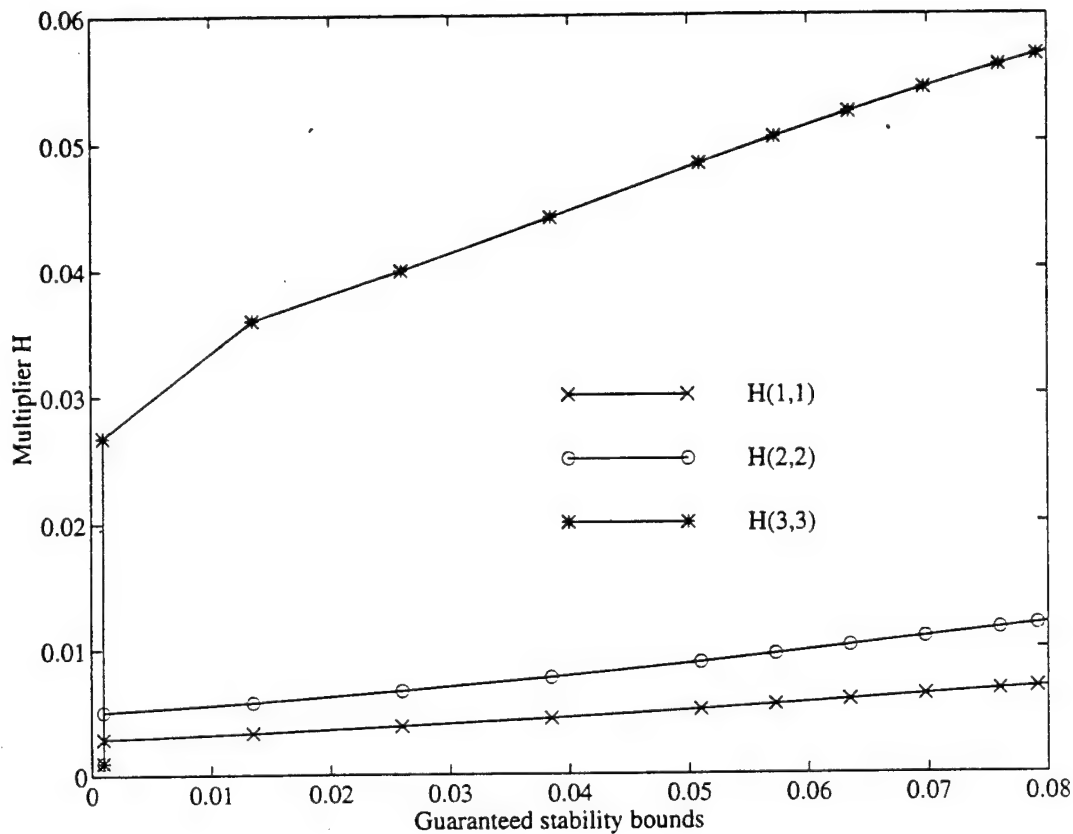
Figure 5: Diagonal elements of the stability multiplier H for the analysis performed using the ME controller.

(a factor of 10) corresponds to a significant performance improvement. The location of the vertical asymptotes in each of the curves corresponds to the limits of robust stability for that particular compensator.

For each controller, the symbol "$\times$" on the performance axis indicates the $\mathcal{H}_2$ performance achieved by that design on the nominal design model. In each case, it can be seen that the nominal performance and the worst case overbound are quite close. This observation agrees with the results in Ref. [19], and further indicates that the Popov $\mathcal{H}_2$ cost overbound is quite tight. Also note that each of the SWLQG, ME, and MM designs have been implemented on the test article, and that these designs represent the best performance that could be experimentally achieved using that particular technique [2]. Thus, the compensators were not necessarily designed for the same values of ρ and μ. Furthermore, the best stabilizing LQG design could only obtain approximately 4 dB on the hardware, so the LQG example in the figure is presented just to show how sensitive optimal LQG controllers are to changes in the modal frequencies.

17

DTIC COULD NOT GET MISSING

PAGES FROM CONTRIBUTOR

18

is a correction step which changes the multipliers at the initial value of λ_0. These results show that the multiplier coefficients change significantly with λ, which indicates the need for a good algorithm for predicting the changes from $\eta(\lambda_i)$ to $\eta(\lambda_i + \Delta\lambda)$. The symbols on the graphs correspond to the results at the end of a prediction–correction iteration.

The results of this analysis can also be used to initialize a redesign of the robust controllers using Popov Controller Synthesis [11,19]. An example of this procedure is illustrated in Figure 6. The stable MM controller was used as the initial design, along with the multipliers computed at $M_2(\lambda) = 0.0585$. This point is indicated by the symbol "o" in Figure 6. The controller was redesigned using a synthesis technique that optimizes the cost overbound with respect to both the multiplier coefficients and the controller gains. The synthesis optimization and stopping criterion are similar to the correction step in the analysis algorithm. In Figure 6, the synthesis corresponds to a reduction in the cost overbound at a constant value of $M_2(\lambda)$. The final result is shown in the figure by the symbol "$\times$" at $M_2(\lambda) = 0.0585$. As indicated, the robust performance at this level of the guaranteed stability bounds has been improved by almost 1 dB.

To complete this example, the new Popov design was then analyzed in two different ways. The algorithm was started at $\lambda_0 = 1 \times 10^{-3}$ with the same initial values used in the original MM analysis. The algorithm was also started using the multiplier values calculated by the synthesis code. The fact that these two analysis curves essentially over plot indicates that the analysis procedure is not overly sensitive to the initial conditions. These analysis results indicate that the Popov design achieves superior nominal and robust performance as compared to the MM controller, and that the robust stability boundaries are substantially improved ($\sim$ 20%).

Thus, Figure 3 illustrates the utility of this analysis tool in an iterative control design methodology based on several robustness techniques with differing capabilities and computational requirements. Furthermore, Figure 6 demonstrates that the Popov analysis and synthesis techniques can be combined to overcome the difficulty of developing initial values for the stability multipliers at large guaranteed stability bounds.

6   Conclusions

Good robust performance analysis plays a critical part in the design of robust controllers. Previous results have shown that the Popov criterion is much less conservative than small gain tests for systems with real parameter uncertainties. This paper extends this earlier work by developing algorithms to analyze systems with multiple uncertainties that have both upper and lower sector-bounds. The typical application of this procedure was demonstrated using robust controllers for MACE. The combined optimal analysis and synthesis algorithms were also used to design a new controller that yields better robust $\mathcal{H}_2$ performance with larger

DTIC COULD NOT GET MISSING

PAGES FROM CONTRIBUTOR

20

[12] V. M. Popov, "On absolute stability of non-linear automatic control systems," *Avtomatika I Telemekhanika*, vol. 22, no. 8, pp. 961–979, 1961.

[13] W. M. Haddad and D. S. Bernstein, "The multivariable parabola criterion for robust controller synthesis: A Riccati equation approach," accepted for publication in the *J. Math. Syst. Est. Contr.*, 1993.

[14] J. P. How and S. R. Hall, "Connections between absolute stability theory and bounds for the structured singular value," accepted for publication in the *IEEE Trans. on Auto. Control*, Dec. 1993. MIT SERC Report #9-92-J, May, 1992.

[15] W. M. Haddad and D. S. Bernstein, "Parameter-dependent Lyapunov functions, constant real parameter uncertainty, and the Popov criterion in robust analysis and synthesis, Parts I and II," in *IEEE Conference on Decision and Control*, pp. 2274–2279, 2632–2633, Dec. 1991. Submitted to the *IEEE Transactions on Automatic Control*.

[16] J. P. How and S. R. Hall, "Connections between the Popov stability criterion and bounds for real parameter uncertainty," in *American Control Conference*, pp. 1084–1089, Inst. of Electrical and Electronics Engineers, Piscataway, NJ, June 1993.

[17] K. S. Narendra and J. H. Taylor, *Frequency Domain Criteria for Absolute Stability*. New York: Academic Press, 1973.

[18] W. M. Haddad, J. P. How, S. R. Hall, and D. S. Bernstein, "Extensions of mixed-μ bounds to monotonic and odd monotonic nonlinearities using absolute stability theory," in *IEEE Conference on Decision and Control*, pp. 2813–2819, 2820–2823, December 1992. Accepted for publication in the *International Journal of Control*.

[19] J. P. How, W. M. Haddad, and S. R. Hall, "Robust control synthesis examples with real parameter uncertainty using the Popov criterion," in *American Control Conference*, pp. 1090–1095, Inst. of Electrical and Electronics Engineers, Piscataway, NJ, June 1993. Accepted for publication in the *AIAA Journal of Guidance, Control, and Dynamics*.

[20] J. C. Hsu and A. U. Meyer, *Modern Control Principles and Applications*. New York: McGraw-Hill Book Company, 1968.

[21] E. G. Collins, Jr., W. M. Haddad, and L. D. Davis, "Riccati equation approaches for robust stability and performance analysis using the small gain, positivity, and Popov Theorems," in *American Control Conference*, pp. 1079–1083, Inst. of Electrical and Electronics Engineers, Piscataway, NJ, 1993. accepted for publication in the *AIAA Journal of Guidance, Control, and Dynamics*.

[22] R. Glaese and D. W. Miller, "On-orbit modelling of the middeck active control experiment from 1-g analysis and experimentation," in to appear at the *12th International Modal Analysis Conference (IMAC)*, Feb. 1994.

[23] J. P. How, S. R. Hall, and W. M. Haddad, "Robust controllers for the Middeck Active Control Experiment using Popov controller synthesis," in *AIAA Guidance, Navigation, and Control Conference*, pp. 1611–1621, Aug. 1993. *Accepted for publication in the* IEEE Transactions on Control Systems Technology.

[24] W. M. Haddad and D. S. Bernstein, "Explicit construction of quadratic Lyapunov functions for the small gain, positivity, circle, and Popov theorems and their application to robust stability part I and II:," *Int. J. Robust and Nonlinear Control, to appear*, 1993.

[25] P. M. Young and J. C. Doyle, "Properties of the mixed-μ problem and its bounds," Preprint, submitted to *IEEE Transactions on Automatic Control*, Oct. 1992.

[26] E. G. Collins, Jr., L. D. Davis, and S. Richter, "Design of reduced-order, $\mathcal{H}_2$ optimal controllers using a homotopy algorithm," in *American Control Conference*, pp. 2658–

DTIC COULD NOT GET MISSING

PAGES FROM CONTRIBUTOR

22

Appendix: Jacobian Expressions

As discussed in Section 4, solving the homotopy problem requires the computation of the Jacobian of the homotopy map. The first part of the Jacobian in Eq. 36 is given by the symmetric matrix

$$
L_\eta = \begin{bmatrix}
\dfrac{\partial H_1}{\partial \alpha} & \cdots & \dfrac{\partial H_1}{\partial N_{ij}} & \cdots & \dfrac{\partial H_1}{\partial H_{ij}} \\[2ex]
\bullet & \cdots \text{vec}_N(\dfrac{\partial H_2}{\partial N_{ij}}) & \cdots & \text{vec}_N(\dfrac{\partial H_2}{\partial H_{ij}}) \\[2ex]
\bullet & \cdots & \bullet & \cdots \text{vec}_N(\dfrac{\partial H_3}{\partial H_{ij}})
\end{bmatrix},
\tag{48}
$$

where H_1, H_2, and H_3 are defined in Eqs. 25–27. Note that only the free parameters of N and H are considered in the calculation of L_η. Also, the columns of L_η are arranged to be consist with the result produced by the $\text{vec}_N(\cdot)$ operator. Let $E_{ij} = e_i e_j^T$, where each element of the column e_i is zero, except for the ith term which is unity. In the following, the dimension of E_{ij} is the same as the dimension of H. For convenience, we first note that

$$
\frac{\partial \overline{C}}{\partial N_{ij}} = E_{ij} C_0 A_m,
\tag{49}
$$

$$
\frac{\partial \overline{C}}{\partial H_{ij}} = E_{ij} C_0,
\tag{50}
$$

$$
\frac{\partial C_{\text{ricc}}}{\partial \theta_j} = \frac{\partial \overline{C}}{\partial \theta_j} - B_0^T \frac{\partial P}{\partial \theta_j},
\tag{51}
$$

$$
\frac{\partial A_1}{\partial \theta_j} = -B_0 \left(\frac{\partial R_0^{-1}}{\partial \theta_j} \overline{C} + R_0^{-1} \frac{\partial \overline{C}}{\partial \theta_j} \right),
\tag{52}
$$

$$
\frac{\partial R_0^{-1}}{\partial N_{ij}} = -R_0^{-1} (E_{ij} C_0 B_0 + (E_{ij} C_0 B_0)^T) R_0^{-1},
\tag{53}
$$

$$
\frac{\partial R_0^{-1}}{\partial H_{ij}} = -R_0^{-1} (E_{ij} (M_2 - M_1)^{-1} + (E_{ij} (M_2 - M_1)^{-1})^T) R_0^{-1},
\tag{54}
$$

$$
\frac{\partial \overline{C}^T R_0^{-1} \overline{C}}{\partial \theta_j} = \overline{C}^T \frac{\partial R_0^{-1}}{\partial \theta_j} \overline{C} + \overline{C}^T R_0^{-1} \frac{\partial \overline{C}}{\partial \theta_j} + \left(\overline{C}^T R_0^{-1} \frac{\partial \overline{C}}{\partial \theta_j} \right)^T,
\tag{55}
$$

where, as before, θ_j refers to any free element of H or N. The gradient expressions, in turn, depend on the solutions to the equations

$$
0 = A_2^T \frac{\partial P}{\partial \alpha} + \frac{\partial P}{\partial \alpha} A_2 + R_{xx},
\tag{56}
$$

$$
0 = A_2 \frac{\partial Q}{\partial \alpha} + \frac{\partial Q}{\partial \alpha} A_2^T + B_0 R_0^{-1} B_0^T \frac{\partial P}{\partial \alpha} Q + \left(B_0 R_0^{-1} B_0^T \frac{\partial P}{\partial \alpha} Q \right)^T - \frac{1}{\alpha^2} V_{xx},
\tag{57}
$$

23

DTIC COULD NOT GET MISSING

PAGES FROM CONTRIBUTOR

24

$$\frac{\partial C_{\mathrm{ricc}}}{\partial \lambda} = \frac{\partial \overline{C}}{\partial \lambda} - B_0^T \frac{\partial P}{\partial \lambda}, \tag{68}$$

$$\frac{\partial A_1}{\partial \lambda} = -B_0[(M_{1f} - M_{1o})C_0 + \frac{\partial R_0^{-1}}{\partial \lambda}\overline{C} + R_0^{-1}\frac{\partial \overline{C}}{\partial \lambda}], \tag{69}$$

$$\frac{\partial \overline{C}^T R_0^{-1} \overline{C}}{\partial \lambda} = \overline{C}^T \frac{\partial R_0^{-1}}{\partial \lambda}\overline{C} + \overline{C}^T R_0^{-1}\frac{\partial \overline{C}}{\partial \lambda} + (\overline{C}^T R_0^{-1}\frac{\partial \overline{C}}{\partial \lambda})^T, \tag{70}$$

$$\frac{\partial R_0^{-1}}{\partial \lambda} = R_0^{-1}([H(M_2 - M_1)^{-1}((M_{2f} - M_{2o}) - (M_{1f} - M_{1o}))(M_2 - M_1)^{-1}]$$
$$+ [H(M_2 - M_1)^{-1}((M_{2f} - M_{2o}) - (M_{1f} - M_{1o}))(M_2 - M_1)^{-1}]^T)R_0^{-1}, \tag{71}$$

which are based on the solutions to the equations

$$0 = A_2^T \frac{\partial P}{\partial \lambda} + \frac{\partial P}{\partial \lambda}A_2 + \frac{\partial A_1}{\partial \lambda}^T P + P\frac{\partial A_1}{\partial \lambda} + PB_0\frac{\partial R_0^{-1}}{\partial \lambda}B_0^T P + \frac{\partial \overline{C}^T R_0^{-1} \overline{C}}{\partial \lambda}, \tag{72}$$

$$0 = A_2\frac{\partial Q}{\partial \lambda} + \frac{\partial Q}{\partial \lambda}A_2^T$$
$$+ [(\frac{\partial A_1}{\partial \lambda} + B_0(\frac{\partial R_0^{-1}}{\partial \lambda}B_0^T P + R_0^{-1}B_0^T\frac{\partial P}{\partial \lambda}))Q]$$
$$+ [(\frac{\partial A_1}{\partial \lambda} + B_0(\frac{\partial R_0^{-1}}{\partial \lambda}B_0^T P + R_0^{-1}B_0^T\frac{\partial P}{\partial \lambda}))Q]^T. \tag{73}$$

Then L_λ in Eq.66 can be written in terms of

$$\frac{\partial H_1}{\partial \lambda} = \mathrm{tr}\,[\frac{\partial Q}{\partial \lambda}R_{\mathbf{xx}}] - \frac{1}{\alpha^2}\mathrm{tr}\,\{(\frac{\partial P}{\partial \lambda} + C_0^T[(M_{2f} - M_{2o}) - (M_{1f} - M_{1o})]NC_0)V_{\mathbf{xx}}\}, \tag{74}$$

$$\frac{\partial H_2}{\partial \lambda} = \frac{1}{\alpha}[(M_{2f} - M_{2o}) - (M_{1f} - M_{1o})]^T C_0 V_{\mathbf{xx}} C_0^T$$
$$+ 2[(\frac{\partial R_0^{-1}}{\partial \lambda}C_{\mathrm{ricc}} + R_0^{-1}\frac{\partial C_{\mathrm{ricc}}}{\partial \lambda})QA_2^T + R_0^{-1}C_{\mathrm{ricc}}\{\frac{\partial Q}{\partial \lambda}A_2^T$$
$$- Q(\frac{\partial C_{\mathrm{ricc}}}{\partial \lambda}^T R_0^{-1} + C_{\mathrm{ricc}}^T\frac{\partial R_0^{-1}}{\partial \lambda} + C_0^T(M_{1f} - M_{1o}))B_0^T\}]C_0^T, \tag{75}$$

$$\frac{\partial H_3}{\partial \lambda} = 2[(\frac{\partial R_0^{-1}}{\partial \lambda}C_{\mathrm{ricc}} + R_0^{-1}\frac{\partial C_{\mathrm{ricc}}}{\partial \lambda})Q(C_0 - (M_2 - M_1)^{-1}R_0^{-1}C_{\mathrm{ricc}})^T$$
$$+ R_0^{-1}C_{\mathrm{ricc}}\{\frac{\partial Q}{\partial \lambda}(C_0 - (M_2 - M_1)^{-1}R_0^{-1}C_{\mathrm{ricc}})^T$$
$$- Q[\frac{\partial C_{\mathrm{ricc}}}{\partial \lambda}^T R_0^{-1} + C_{\mathrm{ricc}}^T(\frac{\partial R_0^{-1}}{\partial \lambda} - R_0^{-1}(M_2 - M_1)^{-1}$$
$$[(M_{2f} - M_{2o}) - (M_{1f} - M_{1o})])](M_2 - M_1)^{-1}\}]. \tag{76}$$

Note that each of the Lyapunov equations for the derivative terms has the same dynamics matrix, A_2. In this work, the Lyapunov equations are solved using an eigenvalue decomposition of A_2. Thus, the decomposition need only be performed once per Hessian calculation, which significantly decreases the computational effort required to determine this matrix.

Appendix H:

Generalized Fixed-Structure Optimality Conditions for H_2 Optimal Control

Generalized Fixed-Structure Optimality Conditions
for H₂ Optimal Control

by

Emmanuel G. Collins, Jr.
Harris Corporation
Government Aerospace Systems Division
MS 19/4848
Melbourne, FL 32902

Wassim M. Haddad
School of Aerospace Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Sidney S. Ying
Rockwell International
Collins Commercial Avionics
MS 306-100
Melbourne, FL 32934

Abstract

Over the last several years, researchers have shown that when it is assumed *a priori* that a fixed-order optimal compensator is minimal, the necessary conditions can be characterized in terms of coupled Riccati and Lyapunov equations, usually termed "optimal projection equations." When the optimal projection equations for H_2 optimal control are specialized to full-order control, the standard LQG Riccati equations are recovered. This paper relaxes the minimality assumption on the compensator and derives necessary conditions for fixed-structure H_2 optimal control that reduce to the standard optimal projection equations when the optimal compensators are assumed to be minimal. The results are then specialized to full-order control. The results show that the standard LQG Riccati equations can be derived using fixed-structure theory even without the minimality assumption. They also show for the first time that a reduced-order optimal projection controller is a projection, described by a projection matrix μ, of one of the extremals (a "central extremal") to the full-order H_2 optimal control problem. For nonminimal LQG compensators the projection matrix ν used in balanced controller reduction produces a minimal-order realization of the LQG compensator, which is of course an optimal reduced-order compensator. For this special case, the similarity transformations relating ν, μ and the optimal projection matrix τ from standard optimal projection theory are explicitly defined. Finally, an illustrative numerical example is presented to demonstrate the design framework discussed in this paper for H_2 optimal, reduced-order, dynamic compensation.

Running Title: Fixed-Structure H_2 Optimal Control

Key Words: Optimal Projection Equations, H_2 Optimal Control, Constrained Optimization, Reduced-Order Control.

---

1. Introduction

One of the foundational results in modern control theory is the development of a characterization of the globally optimal H_2 controller via algebraic Riccati equations [1-3]. This result has traditionally been derived via the Calculus of Variations or the Maximum Principle in conjunction with the Separation Principle [2-4]. Unfortunately, the optimal H_2 or LQG (Linear-Quadratic-Gaussian) controller has dimension equal to that of the plant, although it may have minimal dimension which is less than that of the plant. This has motivated the search for optimal reduced-order controllers, that is, controllers that have dimension less than that of the plant.

Because the Calculus of Variations and the Maximum Principle characterize globally optimal solutions, these traditional methods for deriving the LQG result do not extend to the development of characterizations of optimal reduced-order controllers. Hence, researchers have developed optimization methods that allow the dimension and structure of the controller to be constrained *a priori* (see, e.g., [5-9]). These methods are usually based on Lagrange multiplier theory and will be called here "fixed-structure approaches." The "optimal projection" characterization of the necessary conditions for optimal reduced-order control [6] was derived using a fixed structure approach and yields the standard LQG regulator and observer Riccati equations when the dimension of the controller is specified to be equal to the dimension of the plant. However, the original optimal projection results and numerous extensions (e.g., [7-9]) were derived by *a priori* assuming that the controller is minimal. This is a limiting assumption since it is known that even an LQG controller is not always minimal [10]. It should be noted here that the LQG Riccati equations are also derived in [11] using a fixed-structure approach. However, the results there *a priori* assume that $A_c = A + BC_c - B_cC$, where (A, B, C) is the plant triple and (A_c, B_c, C_c) denotes the controller triple.

This paper presents optimality conditions that are derived without assuming the minimality of the compensator. A similar approach was also considered in [12]. The results are specialized to the case in which the compensator is constrained to have the dimension of the plant. It is shown that even when compensator minimality is not assumed, fixed-structure theory is able to derive the LQG Riccati equations. It is also shown that there exist sets of coupled Riccati and Lyapunov equations that are identical in form to the optimal projection equations for reduced-order control but actually characterize extremals to the full-order compensation problem. This leads to a new interpretation of an optimal projection controller. In particular, an optimal projection controller

1

is a projection, described by a projection matrix μ, of a "central" extremal to the H_2 optimal, full-order compensation problem.

For nonminimal LQG compensators the projection matrix ν used in balanced controller reduction produces a minimal-order realization of the LQG compensator, which is of course an optimal reduced-order compensator. For this special case, the similarity transformations relating ν, μ, and the optimal projection matrix τ from the standard optimal projection theory are explicitly defined.

The primary reason for developing the Riccati equation approach to reduced-order dynamic compensation is to enable the development of efficient computational algorithms for controller synthesis. In particular, the goal has been to develop algorithms that exploit the special structure of the Riccati equations. This paper gives a brief overview of the continuation algorithm developed in [13, 14] that utilizes the special Riccati-equation structure. The results are illustrated by developing reduced-order controllers for an important benchmark problem in structural control.

The paper is organized as follows. Section 2 presents the optimal fixed-structure dynamic compensation problem and some preliminary lemmas. Section 3 develops necessary conditions characterizing solutions to the optimal fixed-structure dynamic problem without an *a priori* minimal compensator assumption. Next, Section 4 specializes the optimality conditions to the case of full-order dynamic compensation and discusses the relationship between a "central" extremal and the LQG compensator. Section 5 demonstrates the utility of the H_2 optimal reduced-order controller design framework discussed in the previous sections with a benchmark numerical example. Finally, Section 6 presents the conclusions.

Notation

$\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r$	real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$
$\mathbb{E}$	expected value
$\mathcal{R}(X), \mathcal{N}(X)$	range space of matrix X, null space of matrix X
$X^\dagger$	*Moore-Penrose generalized inverse* of matrix X [15]
$X^\#$	*group inverse* of matrix X [15]
$X \geq 0, X > 0$	matrix X is nonnegative definite, matrix X is positive definite
$0_{r \times s}$	$r \times s$ zero matrix
I_r	$r \times r$ identity matrix

2. The Optimal Fixed-Structure Dynamic Compensation Problem

Consider the nth-order linear time-invariant plant

$$\dot{x}(t) = Ax(t) + Bu(t) + D_1 w(t), \tag{2.1a}$$

$$y(t) = Cx(t) + D_2 w(t), \tag{2.1b}$$

where (A, B) is stabilizable, (A, C) is detectable, $x \in \mathbb{R}^n, u \in \mathbb{R}^m, y \in \mathbb{R}^l$, and $w \in \mathbb{R}^d$ is a standard white noise disturbance with intensity I_d and rank $D_2 = l$. The intensities of $D_1 w(t)$ and $D_2 w(t)$ are thus given, respectively, by $V_1 \triangleq D_1 D_1^T \geq 0$, and $V_2 \triangleq D_2 D_2^T > 0$. For convenience, we assume that $V_{12} \triangleq D_1 D_2^T = 0$, i.e., the plant disturbance and measurement noise are uncorrelated. The goal of the *optimal fixed-order dynamic compensation problem* is to determine an n_c^{th}-order dynamical compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \tag{2.2a}$$

$$u(t) = -C_c x_c(t), \tag{2.2b}$$

which satisfies the following two design criteria:

(i) the closed-loop system corresponding to (2.1) and (2.2) given by

$$\dot{\tilde{x}}(t) = \tilde{A}\tilde{x}(t) + \tilde{D}w(t), \tag{2.3}$$

where

$$\tilde{x}(t) \triangleq \begin{bmatrix} x(t) \\ x_c(t) \end{bmatrix}, \quad \tilde{A} \triangleq \begin{bmatrix} A & -BC_c \\ B_c C & A_c \end{bmatrix}, \quad \tilde{D} \triangleq \begin{bmatrix} D_1 \\ B_c D_2 \end{bmatrix}, \tag{2.4, 5, 6}$$

is asymptotically stable; and

(ii) the compensator minimizes the steady-state quadratic performance criterion

$$J(A_c, B_c, C_c) \triangleq \lim_{t \to \infty} \frac{1}{t} \mathbb{E} \int_0^t [x^T(s) R_1 x(s) + u^T(s) R_2 u(s)] ds, \tag{2.7}$$

where $R_1 \geq 0$ and $R_2 > 0$.

Although a cross-weighting term of the form $2x^T(t) R_{12} u(t)$ can also be included in (2.7), we shall not do so here to facilitate the presentation. With the first criterion, we restrict our attention to the set of stabilizing compensators, $\mathcal{S}_c \triangleq \{(A_c, B_c, C_c): \tilde{A} \text{ is asymptotically stable}\}$

3

which guarantees that the cost J is finite and independent of initial conditions. The cost (2.7) can now be expressed as

$$J(A_c, B_c, C_c) = \lim_{t \to \infty} \mathbb{E}[\tilde{x}^{\mathrm{T}}(t)\tilde{R}\tilde{x}(t)], \tag{2.8}$$

where

$$\tilde{R} \triangleq \begin{bmatrix} R_1 & 0 \\ 0 & C_c^{\mathrm{T}} R_2 C_c \end{bmatrix}. \tag{2.9}$$

Next, by introducing the performance variable

$$z(t) \triangleq E_1 x(t) + E_2 u(t) = \tilde{E}\tilde{x}(t), \tag{2.10}$$

and defining the transfer function from disturbances w to performance variables z by

$$\tilde{H}(s) \triangleq \tilde{E}(sI_{\tilde{n}} - \tilde{A})^{-1} \tilde{D},$$

where $\tilde{E} \triangleq [E_1 \ E_2 C_c]$, and $\tilde{n} \triangleq n + n_c$, it can be shown that when $\tilde{A}$ is asymptotically stable, (2.8) is given by $J(A_c, B_c, C_c) = \|\tilde{H}(s)\|_2^2$. For convenience we thus define the matrices $R_1 \triangleq E_1^{\mathrm{T}} E_1$ and $R_2 \triangleq E_2^{\mathrm{T}} E_2$ which are the H_2 weights for the state and control variables. Since $\tilde{A}$ is asymptotically stable, there exist nonnegative-definite matrices $\tilde{Q} \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ and $\tilde{P} \in \mathbb{R}^{\tilde{n} \times \tilde{n}}$ satisfying the closed-loop steady-state covariance equation and its dual, i.e.,

$$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^{\mathrm{T}} + \tilde{V}, \tag{2.11}$$

$$0 = \tilde{A}^{\mathrm{T}}\tilde{P} + \tilde{P}\tilde{A} + \tilde{R}, \tag{2.12}$$

where

$$\tilde{V} \triangleq \begin{bmatrix} V_1 & 0 \\ 0 & B_c V_2 B_c^{\mathrm{T}} \end{bmatrix}. \tag{2.13}$$

The cost functional (2.7) can now be expressed as

$$J(A_c, B_c, C_c) = \mathrm{tr}\,\tilde{Q}\tilde{R} = \mathrm{tr}\,\tilde{P}\tilde{V}. \tag{2.14}$$

Furthermore, $\tilde{Q}$ and $\tilde{P}$ can be expressed in the partitioned forms

$$\tilde{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^{\mathrm{T}} & Q_2 \end{bmatrix}, Q_1 \in \mathbb{R}^{n \times n}, Q_2 \in \mathbb{R}^{n_c \times n_c}, \tag{2.15}$$

$$\tilde{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^{\mathrm{T}} & P_2 \end{bmatrix}, P_1 \in \mathbb{R}^{n \times n}, P_2 \in \mathbb{R}^{n_c \times n_c}. \tag{2.16}$$

4

Note that Q_1 is the covariance of the plant states, Q_2 is the covariance of the compensator states and Q_{12} is the cross-covariance of the plant and controller states. Using (2.6), (2.9), (2.13) (2.15) and (2.16), and expanding (2.11) and (2.12), yields

$$0 = AQ_1 + Q_1 A^{\mathrm{T}} - BC_c Q_{12}^{\mathrm{T}} - Q_{12} C_c^{\mathrm{T}} B^{\mathrm{T}} + V_1, \tag{2.17}$$

$$0 = AQ_{12} + Q_{12} A_c^{\mathrm{T}} - BC_c Q_2 + Q_1 C^{\mathrm{T}} B_c^{\mathrm{T}}, \tag{2.18}$$

$$0 = A_c Q_2 + Q_2 A_c^{\mathrm{T}} + B_c C Q_{12} + Q_{12}^{\mathrm{T}} C^{\mathrm{T}} B_c^{\mathrm{T}} + B_c V_2 B_c^{\mathrm{T}}, \tag{2.19}$$

$$0 = A^{\mathrm{T}} P_1 + P_1 A + C^{\mathrm{T}} B_c^{\mathrm{T}} P_{12}^{\mathrm{T}} + P_{12} B_c C + R_1, \tag{2.20}$$

$$0 = A^{\mathrm{T}} P_{12} + P_{12} A_c + C^{\mathrm{T}} B_c^{\mathrm{T}} P_2 - P_1 BC_c, \tag{2.21}$$

$$0 = A_c^{\mathrm{T}} P_2 + P_2 A_c - P_{12}^{\mathrm{T}} BC_c - C_c^{\mathrm{T}} B^{\mathrm{T}} P_{12} + C_c^{\mathrm{T}} R_2 C_c. \tag{2.22}$$

Before presenting the main theorems we present the following key lemmas and definitions which are useful in stating and proving the main results of the paper. First, we introduce the notion of a projection. Specifically, let $\mathcal{X}_1$ and $\mathcal{X}_2$ denote subspaces of a linear vector space $\mathcal{X}$. Then $\mathcal{X}_1$ and $\mathcal{X}_2$ are a *decomposition* of $\mathcal{X}$ if $\mathcal{X}_1$ and $\mathcal{X}_2$ are disjoint, i.e., $\mathcal{X}_1 \cap \mathcal{X}_2 = 0$, and every vector $x \in \mathcal{X}$ can be uniquely expressed as $x = x_1 + x_2$, such that $x_1 \in \mathcal{X}_1$ and $x_2 \in \mathcal{X}_2$. $\mathcal{X}$ is called the *direct sum* of $\mathcal{X}_1$ and $\mathcal{X}_2$, or, $\mathcal{X} = \mathcal{X}_1 \oplus \mathcal{X}_2$.

Definition 2.1. (Rao and Mitra [15]). The linear operator $L : x \to x_1$ is called a *projection* on $\mathcal{X}_1$ along $\mathcal{X}_2$.

Definition 2.2. A $n \times n$ matrix τ is a *projection matrix* if τ is idempotent, i.e., $\tau^2 = \tau$.

Remark 2.1. A projection matrix τ defines a projection on $\mathcal{R}(\tau)$ along $\mathcal{N}(\tau)$, which implies that for two projection matrices τ_1 and τ_2, if $\mathcal{R}(\tau_1) = \mathcal{R}(\tau_2)$ and $\mathcal{N}(\tau_1) = \mathcal{N}(\tau_2)$ then τ_1 and τ_2 define the same projection.

Lemma 2.1. (Rao and Mitra [15]). Suppose X and Y are matrices with compatible dimensions. Then

(*i*) $\mathcal{N}(Y) \subset \mathcal{N}(X)$ if and only if $XY^{\dagger}Y = X$.

(*ii*) $\mathcal{R}(Y) \subset \mathcal{R}(X)$ if and only if $XX^{\dagger}Y = Y$.

Lemma 2.2. Assume $\tilde{A}$ is asymptotically stable and $\tilde{Q}$ and $\tilde{P}$ are the $\tilde{n} \times \tilde{n}$ nonnegative-definite solutions of (2.11) and (2.12). Let $\tilde{Q}$ and $\tilde{P}$ be partitioned as in (2.15) and (2.16). Then, the following relations hold:

5

(i) $\mathcal{N}(P_2) \subset \mathcal{N}(P_{12})$.

(ii) $\mathcal{N}(P_2) \subset \mathcal{N}(C_c)$.

(iii) $\mathcal{N}(P_2)$ is the unobservable subspace of (A_c, C_c).

(iv) $\mathcal{N}(Q_2) \subset \mathcal{N}(Q_{12})$.

(v) $\mathcal{N}(Q_2) \subset \mathcal{N}(B_c^{\mathrm{T}})$.

(vi) $\mathcal{N}(Q_2)$ is the uncontrollable subspace of (A_c, B_c).

Proof. Let $v \in \mathcal{N}(P_2)$, or, equivalently, $P_2 v = 0$. Consider a vector x defined by

$$x \triangleq \begin{bmatrix} P_{12} v \\ \alpha v \end{bmatrix},$$

where $\alpha \in \mathbb{R}$ is arbitrary. It follows from $\tilde{P} \geq 0$ that

$$x^{\mathrm{T}} \tilde{P} x = v^{\mathrm{T}} P_{12}^{\mathrm{T}} P_1 P_{12} v + 2 \alpha v^{\mathrm{T}} P_{12}^{\mathrm{T}} P_{12} v \geq 0.$$

The above expression is true for all $\alpha \in \mathbb{R}$ if and only if $P_{12} v = 0$, which implies $v \in \mathcal{N}(P_{12})$, or, equivalently, $\mathcal{N}(P_2) \subset \mathcal{N}(P_{12})$.

Forming $v^{\mathrm{T}}(2.22)v$ and noting $P_2 v = 0$ and $P_{12} v = 0$, yields $v^{\mathrm{T}} C_c^{\mathrm{T}} R_2 C_c v = 0$. Since $R_2 > 0$, we obtain $C_c v = 0$, which implies $v \in \mathcal{N}(C_c)$, or, equivalently, $\mathcal{N}(P_2) \subset \mathcal{N}(C_c)$.

Forming $(2.22)v$ and noting $C_c v = 0$, $P_2 v = 0$, and $P_{12} v = 0$, yields $P_2 A_c v = 0$. Using property (ii), $P_2 A_c v = 0$ implies $C_c A_c v = 0$. Thus, $v \in$ unobservable subspace of (A_c, C_c), or, equivalently, $\mathcal{N}(P_2) \subset$ unobservable subspace of (A_c, C_c). Using the dual approach, properties (iv),(v) and (vi) can be verified. $\square$

Lemma 2.3. (Albert [16]). If $\tilde{A}$ is asymptotically stable, then

$$P_{12} = P_{12} P_2^\dagger P_2, \quad P_{12}^{\mathrm{T}} = P_2 P_2^\dagger P_{12}^{\mathrm{T}}, \tag{2.23a,b}$$

$$Q_{12} = Q_{12} Q_2^\dagger Q_2, \quad Q_{12}^{\mathrm{T}} = Q_2 Q_2^\dagger Q_{12}^{\mathrm{T}}, \tag{2.24a,b}$$

$$C_c = C_c P_2^\dagger P_2, \quad C_c^{\mathrm{T}} = P_2 P_2^\dagger C_c^{\mathrm{T}}, \tag{2.25a,b}$$

$$B_c = Q_2 Q_2^\dagger B_c, \quad B_c^{\mathrm{T}} = B_c^{\mathrm{T}} Q_2^\dagger Q_2. \tag{2.26a,b}$$

Proof. The result is a direct consequence of Lemma 2.1 along with statements (i), (ii),(iv), and (v) of Lemma 2.2. $\square$

6

Lemma 2.4. Suppose $\hat{Q} \in \mathbb{R}^{n \times n}$ and $\hat{P} \in \mathbb{R}^{n \times n}$ are nonnegative definite with rank $\hat{Q} = n_q$, rank $\hat{P} = n_p$ and rank $\hat{Q}\hat{P} = n_r$. Then, the following statements hold:

(i) There exists invertible $W \in \mathbb{R}^{n \times n}$ such that $W^{-1}\hat{Q}W^{-T}$ and $W^T\hat{P}W$ are both diagonal.

(ii) $\hat{Q}\hat{P}$ is diagonalizable and has nonnegative eigenvalues.

(iii) The $n \times n$ matrix

$$\tau \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#} \tag{2.27}$$

is idempotent, i.e., τ is a projection matrix and rank $\tau = n_r$.

Furthermore, there exists a nonsingular matrix $W \in \mathbb{R}^{n \times n}$ such that

$$\tau = W \begin{bmatrix} I_{n_r} & 0 \\ 0 & 0 \end{bmatrix} W^{-1}. \tag{2.28}$$

In addition, if we define $\tau_\perp \triangleq I_n - \tau$, then rank $\tau_\perp = n - n_r$.

(iv) There exists a nonsingular transformation $W \in \mathbb{R}^{n \times n}$ such that

$$\hat{Q} = W \begin{bmatrix} \Omega_1 & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & \Omega_3 & 0 \\ 0_{n_t \times n_r} & 0 & 0 & 0 \end{bmatrix} W^T, \tag{2.29}$$

$$\hat{P} = W^{-T} \begin{bmatrix} \Omega_1 & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & \Omega_2 & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{n_t \times n_r} & 0 & 0 & 0 \end{bmatrix} W^{-1}, \tag{2.30}$$

where $n_t = n - (n_p + n_q - n_r)$ and $\Omega_1 \in \mathbb{R}^{n_r \times n_r}$, $\Omega_2 \in \mathbb{R}^{(n_p - n_r) \times (n_p - n_r)}$ and $\Omega_3 \in \mathbb{R}^{(n_q - n_r) \times (n_q - n_r)}$ are diagonal and positive definite.

(v) Suppose $n_c \geq n_p + n_q - n_r$. Then there exist $G, \Gamma \in \mathbb{R}^{n_c \times n}$, and $M \in \mathbb{R}^{n_c \times n_c}$ such that

$$\mathcal{R}(G^T) = \mathcal{R}(\hat{Q}), \quad \mathcal{R}(\Gamma^T) = \mathcal{R}(\hat{P}), \text{ and rank } M = \text{rank } \hat{Q}\hat{P} = n_r \tag{2.31a, b, c}$$

$$\hat{Q}\hat{P} = G^T M \Gamma, \tag{2.32}$$

$$\Gamma G^T = T \begin{bmatrix} I_{n_r} & 0_{n_r \times (n_c - n_r)} \\ 0_{(n_c - n_r) \times n_r} & 0 \end{bmatrix} T^{-1}, \tag{2.33}$$

where T is an arbitrary $n_c \times n_c$ nonsingular matrix.

7

(*vi*) The matrices $G, \Gamma \in \mathbb{R}^{n_c \times n}$ and $M \in \mathbb{R}^{n_c \times n_c}$ satisfying (2.31)-(2.33) are unique except for a change of basis in $\mathbb{R}^{n_c}$, i.e., if G', Γ' and M' also satisfy property (*v*), then there exists nonsingular $S \in \mathbb{R}^{n_c \times n_c}$ such that $G' = S^T G, \Gamma' = S^{-1}\Gamma$, and $M' = S^{-1}MS$.

(*vii*) If G, Γ and M are as in (*v*), then

$$(\hat{Q}\hat{P})^{\#} = G^T M^{\dagger}\Gamma, \tag{2.34}$$

$$\tau \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#} = G^T\Gamma. \tag{2.35}$$

Proof. Properties (*i*)-(*iii*) are stated and proved in [17]. Property (*iv*) is a direct consequence of Theorem 4.3 in [18].

To prove (*v*) we note that using property (*iv*), $\hat{Q}$ and $\hat{P}$ can be contragrediently diagonalized as in (2.29) and (2.30), respectively. Thus, it follows that

$$\hat{Q}\hat{P} = W \begin{bmatrix} \Omega_1^2 & 0_{n_r \times (n - n_r)} \\ 0_{(n - n_r) \times n_r} & 0 \end{bmatrix} W^{-1}.$$

Furthermore, using the fact that for arbitrary dimensionally-compatible matrices X and Y, $\mathcal{R}(X) = \mathcal{R}(XX^T)$ and $\mathcal{R}(XY) = X\mathcal{R}(Y)$, it follows that

$$\mathcal{R}(\hat{Q}) = W\mathcal{R}\left(\begin{bmatrix} \Omega_1^{\frac{1}{2}} & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & \Omega_3^{\frac{1}{2}} & 0 \\ 0_{n_t \times n_r} & 0 & 0 & 0 \end{bmatrix} \right).$$

Similarly, we can obtain

$$\mathcal{R}(\hat{P}) = W^{-T}\mathcal{R}\left(\begin{bmatrix} \Omega_1^{\frac{1}{2}} & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & \Omega_2^{\frac{1}{2}} & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{n_t \times n_r} & 0 & 0 & 0 \end{bmatrix} \right)$$

Next, choose

$$G' = \begin{bmatrix} I_{n_r} & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & I_{n_q - n_r} & 0 \\ 0_{(n_c - (n_p + n_q - n_r)) \times n_r} & 0 & 0 & 0 \end{bmatrix} W^T,$$

$$\Gamma' = \begin{bmatrix} I_{n_r} & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & I_{n_p - n_r} & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{(n_c - (n_p + n_q - n_r)) \times n_r} & 0 & 0 & 0 \end{bmatrix} W^{-1},$$

8

$$M' = \begin{bmatrix} \Omega_1^2 & 0_{n_r \times (n_c - n_r)} \\ 0_{(n_c - n_r) \times n_r} & 0 \end{bmatrix}.$$

Then it is easy to show by construction that (G', M', Γ') satisfy (2.31)-(2.33) with $T = I_{n_c}$ for this particular case which proves property (v).

To prove property (vi), consider a general triple (G, M, Γ) satisfying (2.31)-(2.33). Noting the above expression for $\mathcal{R}(\hat{Q})$, it follows that the general expression for a matrix G satisfying $\mathcal{R}(G^{\mathrm{T}}) = \mathcal{R}(\hat{Q})$ is

$$G = \begin{bmatrix} G_1 & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & G_3 & 0 \\ 0_{(n_c - (n_p + n_q - n_r)) \times n_r} & 0 & 0 & 0 \end{bmatrix} W^{\mathrm{T}},$$

where $G_1 \in \mathrm{IR}^{n_r \times n_r}$ and $G_3 \in \mathrm{IR}^{(n_q - n_r) \times (n_q - n_r)}$ are nonsingular. Noting the structure of G and G', we can always find a nonsingular $n_c \times n_c$ matrix T_G such that $G = T_G^{-1} G'$. Similarly, using the identity that $\mathcal{R}(\Gamma^{\mathrm{T}}) = \mathcal{R}(\hat{P})$, there exists a nonsingular $n_c \times n_c$ matrix T_Γ such that $\Gamma = T_\Gamma \Gamma'$. However, using (2.33) yields $T_\Gamma = T_G = T$. Furthermore, it follows from (2.32) that $M = T M' T^{-1}$. Thus (G, M, Γ) is unique except for a change of basis in IR^{n_c}.

Finally, to prove (vii) it follows from properties (v) and (vi) that for a general (G, M, Γ) satisfying (2.31)-(2.33),

$$G^{\mathrm{T}} M^\dagger \Gamma = W \begin{bmatrix} \Omega_1^{-1} & 0_{n_r \times (n - n_r)} \\ 0_{(n - n_r) \times n_r} & 0 \end{bmatrix} W^{-1} = (\hat{Q}\hat{P})^\#,$$

and $G^{\mathrm{T}} \Gamma = W \begin{bmatrix} I_{n_r} & 0 \\ 0 & 0 \end{bmatrix} W^{-1} = \tau$. $\quad \square$

Definition 2.3. A triple (G, M, Γ) satisfying (2.31)-(2.33) with $G, \Gamma \in \mathrm{IR}^{n_c \times n}$, $M \in \mathrm{IR}^{n_c \times n_c}$, and $n_c \geq n_r = \mathrm{rank}\, \hat{Q}\hat{P}$ is called a n_c^{th}-order generalized projective factorization of $\hat{Q}\hat{P}$. If $n_c = n_r$, then (G, M, Γ) is called a projective factorization of $\hat{Q}\hat{P}$ [17].

Lemma 2.5. Assume that $\tilde{A}$ is asymptotically stable and that $\hat{Q}$ and $\hat{P}$ are the $\bar{n} \times \bar{n}$ nonnegative-definite solutions of (2.11) and (2.12). Let $\tilde{Q}$ and $\tilde{P}$ be partitioned as in (2.15) and (2.16) and define

$$\hat{Q} \triangleq Q_{12} Q_2^\dagger Q_{12}^{\mathrm{T}}, \tag{2.36}$$

$$\hat{P} \triangleq P_{12} P_2^\dagger P_{12}^{\mathrm{T}}. \tag{2.37}$$

Furthermore, let G, M and Γ be given by

$$G = Q_2^\dagger Q_{12}^{\mathrm{T}}, \tag{2.38}$$

9

$$M = Q_2 P_2, \tag{2.39}$$

and

$$\Gamma = -P_2^\dagger P_{12}^{\mathrm{T}}. \tag{2.40}$$

If

$$P_{12}^{\mathrm{T}} Q_{12} + P_2 Q_2 = 0, \tag{2.41}$$

then the following hold:

(i) The nonzero eigenvalues of $\hat{Q}\hat{P}$ and $Q_2 P_2$ are identical.

(ii) The triple (G, M, Γ) satisfies property (v) of Lemma 2.4, or, equivalently, (G, M, Γ) is a n_c^{th}-order generalized projective factorization of $\hat{Q}\hat{P}$.

Proof. First, note that it follows from (2.36) and (2.37) that

$$\hat{Q}\hat{P} = Q_{12} Q_2^\dagger Q_{12}^{\mathrm{T}} P_{12} P_2^\dagger P_{12}^{\mathrm{T}}. \tag{2.42}$$

Next, using (2.41)$^{\mathrm{T}}$, it follows that $Q_{12}^{\mathrm{T}} P_{12} = -Q_2 P_2$ which substituting into (2.42) yields $\hat{Q}\hat{P} = -Q_{12} Q_2^\dagger Q_2 P_2 P_2^\dagger P_{12}^{\mathrm{T}}$. Since $\tilde{A}$ is assumed to be asymptotically stable, using (2.23b) and (2.24a) of Lemma 2.3, yields

$$\hat{Q}\hat{P} = -Q_{12} P_{12}^{\mathrm{T}}. \tag{2.43}$$

Finally, it follows from (2.43) and (2.41)$^{\mathrm{T}}$ that

$$\lambda_i(\hat{Q}\hat{P}) = \lambda_i(-Q_{12} P_{12}^{\mathrm{T}}) = \lambda_i(-P_{12}^{\mathrm{T}} Q_{12}) = \lambda_i(P_2 Q_2),$$

where $\lambda_i(\cdot)$ represents a nonzero eigenvalue of $\hat{Q}\hat{P}$. Thus, the nonzero eigenvalues of $\hat{Q}\hat{P}$ and $Q_2 P_2$ are identical.

Next, without loss of generality, let rank $Q_2 = n_q$ and rank $P_2 = n_p$. Using the property (i), yields

$$\text{rank } Q_2 P_2 = \text{rank } \hat{Q}\hat{P} = n_r. \tag{2.44}$$

Thus, using (2.44) and (2.39) it follows that rank $M = $ rank $\hat{Q}\hat{P}$. Furthermore, it follows from (2.38), (2.39) and (2.40) that

$$G^{\mathrm{T}} M \Gamma = -Q_{12} Q_2^\dagger Q_2 P_2 P_2^\dagger P_{12}^{\mathrm{T}}.$$

10

Next using $(2.41)^T$, yields (2.32). Computing $(2.40)(2.38)^T$, yields $\Gamma G^T = -P_2^\dagger P_{12}^T Q_{12} Q_2^\dagger$. Now using (2.41), we obtain

$$\Gamma G^T = P_2^\dagger P_2 Q_2 Q_2^\dagger. \tag{2.45}$$

Since Q_2 and P_2 are nonnegative definite, using property (iv) of Lemma 2.4 with $n = n_c$, there exists a nonsingular $T \in \mathbb{R}^{n_c \times n_c}$ such that

$$Q_2 = T \begin{bmatrix} \Omega_1 & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & \Omega_3 & 0 \\ 0_{n_t \times n_r} & 0 & 0 & 0 \end{bmatrix} T^T, \tag{2.46}$$

$$P_2 = T^{-T} \begin{bmatrix} \Omega_1 & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & \Omega_2 & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{n_t \times n_r} & 0 & 0 & 0 \end{bmatrix} T^{-1}, \tag{2.47}$$

where $n_t = n_c - (n_p + n_q - n_r)$ and $\Omega_1 \in \mathbb{R}^{n_r \times n_r}$, $\Omega_2 \in \mathbb{R}^{(n_p - n_r) \times (n_p - n_r)}$ and $\Omega_3 \in \mathbb{R}^{(n_q - n_r) \times (n_q - n_r)}$ are diagonal and positive definite.

Forming $(2.46)(2.46)^\dagger$, yields

$$Q_2 Q_2^\dagger = T \begin{bmatrix} I_{n_r} & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & I_{n_q - n_r} & 0 \\ 0_{n_t \times n_r} & 0 & 0 & 0 \end{bmatrix} T^{-1}, \tag{2.48}$$

Similarly, $(2.47)^\dagger (2.47)$ yields

$$P_2^\dagger P_2 = T \begin{bmatrix} I_{n_r} & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & I_{n_p - n_r} & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{n_t \times n_r} & 0 & 0 & 0 \end{bmatrix} T^{-1}, \tag{2.49}$$

Substituting (2.48) and (2.49) into (2.45), yields (2.33). Note that Q_2 is symmetric which implies that $Q_2^\dagger$ is symmetric, $Q_2^\dagger = (Q_2^\dagger)^{\frac{1}{2}} (Q_2^\dagger)^{\frac{1}{2}}$, and $\mathcal{R}(Q_2^\dagger) = \mathcal{R}((Q_2^\dagger)^{\frac{1}{2}})$. Thus, it follows from the basic matrix geometric properties that

$$
\begin{aligned}
\mathcal{R}(\hat{Q}) &= \mathcal{R}(Q_{12} Q_2^\dagger Q_{12}^T) \\
&= \mathcal{R}(Q_{12} (Q_2^\dagger)^{\frac{1}{2}} (Q_2^\dagger)^{\frac{1}{2}} Q_{12}^T) \\
&= \mathcal{R}(Q_{12} (Q_2^\dagger)^{\frac{1}{2}}) \\
&= Q_{12} \mathcal{R}((Q_2^\dagger)^{\frac{1}{2}}) \\
&= Q_{12} \mathcal{R}(Q_2^\dagger) \\
&= \mathcal{R}(Q_{12} Q_2^\dagger) \\
&= \mathcal{R}(G^T).
\end{aligned}
$$

Following the dual approach yields $\mathcal{R}(\hat{P}) = \mathcal{R}(\Gamma^{\mathrm{T}})$. $\square$

With the above collection of lemmas and definitions we proceed in providing necessary conditions for optimality for the generalized fixed-structure dynamic compensation problem. These conditions are developed in the following section.

3. Optimality Conditions for Fixed-Order Dynamic Compensation

In this section we obtain necessary conditions that characterize solutions to the optimal fixed-structure dynamic compensation problem. Unlike previous results, the compensators are allowed to be nonminimal. We begin by presenting the following key definitions.

Definition 3.1. A compensator (A_c, B_c, C_c) is an *extremal* of the optimal generalized fixed-order dynamic compensation problem if it satisfies the first order necessary conditions of optimality, i.e.,

$$\frac{\partial J}{\partial A_c} = 0, \quad \frac{\partial J}{\partial B_c} = 0, \quad \frac{\partial J}{\partial C_c} = 0,$$

where $J(A_c, B_c, C_c)$ is defined in (2.7).

Definition 3.2. A compensator (A_c, B_c, C_c) is an *admissible extremal* of the optimal generalized fixed-order dynamic compensation problem if it is an extremal and is also in $\mathcal{S}_c$, i.e., the closed-loop system is asymptotically stable.

Finally, for convenience in stating the main results we define

$$\bar{\Sigma} \triangleq C^{\mathrm{T}} V_2^{-1} C, \quad \Sigma \triangleq B R_2^{-1} B^{\mathrm{T}}.$$

Theorem 3.1. Suppose (A_c, B_c, C_c) is an admissible extremal of the optimal fixed-order dynamic compensation problem. Then, there exist $n \times n$ nonnegative-definite matrices $P, Q, \hat{P}$ and $\hat{Q}$ such that A_c, B_c and C_c are given by

$$A_c = \Gamma(A - \Sigma P - Q\bar{\Sigma})G^{\mathrm{T}} + Z - P_2^{\dagger} P_2 Z Q_2 Q_2^{\dagger}, \tag{3.1a}$$

$$B_c = \Gamma Q C^{\mathrm{T}} V_2^{-1} + (I_{n_c} - P_2^{\dagger} P_2) X, \tag{3.1b}$$

$$C_c = R_2^{-1} B^{\mathrm{T}} P G^{\mathrm{T}} + Y(I_{n_c} - Q_2 Q_2^{\dagger}), \tag{3.1c}$$

where (G, M, Γ) is a generalized projective factorization of $\hat{Q}\hat{P}$, and $X \in \mathbb{R}^{n_c \times l}, Y \in \mathbb{R}^{m \times n_c}$ and $Z \in \mathbb{R}^{n_c \times n_c}$ in (3.1) are chosen to satisfy the following constraints:

$$P_2 A_c (I_{n_c} - P_2^{\dagger} P_2) = 0, \tag{3.2}$$

12

$$Q_2 A_c^\mathrm{T}(I_{n_c} - Q_2 Q_2^\dagger) = 0, \qquad (3.3)$$

$$\mathcal{N}(P_2) \subset \mathcal{N}(C_c), \qquad (3.4)$$

$$\mathcal{N}(Q_2) \subset \mathcal{N}(B_c^\mathrm{T}). \qquad (3.5)$$

Here, $Q_2 \in \mathrm{IR}^{n_c \times n_c}$ and $P_2 \in \mathrm{IR}^{n_c \times n_c}$ are respectively the closed-loop covariance of the controller states and its dual and $P, Q, \hat{P}$ and $\hat{Q}$ satisfy:

$$0 = A^\mathrm{T} P + PA + R_1 - P\Sigma P + (R_2^{-1} B^\mathrm{T} P - C_c \Gamma)^\mathrm{T} R_2 (R_2^{-1} B^\mathrm{T} P - C_c \Gamma), \qquad (3.6)$$

$$0 = AQ + QA^\mathrm{T} + V_1 - Q\bar{\Sigma}Q + (QC^\mathrm{T} V_2^{-1} - G^\mathrm{T} B_c) V_2 (QC^\mathrm{T} V_2^{-1} - G^\mathrm{T} B_c)^\mathrm{T}, \qquad (3.7)$$

$$0 = (A - Q\bar{\Sigma})^\mathrm{T} \hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P - (R_2^{-1} B^\mathrm{T} P - C_c \Gamma)^\mathrm{T} R_2 (R_2^{-1} B^\mathrm{T} P - C_c \Gamma), \qquad (3.8)$$

$$0 = (A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^\mathrm{T} + Q\bar{\Sigma}Q - (QC^\mathrm{T} V_2^{-1} - G^\mathrm{T} B_c) V_2 (QC^\mathrm{T} V_2^{-1} - G^\mathrm{T} B_c)^\mathrm{T}. \qquad (3.9)$$

Furthermore, the minimal cost is given by

$$J(A_c, B_c, C_c) = \mathrm{tr}[(P + \hat{P})V_1 + \hat{P}Q\bar{\Sigma}Q], \qquad (3.10)$$

or, equivalently,

$$J(A_c, B_c, C_c) = \mathrm{tr}[(Q + \hat{Q})R_1 + \hat{Q}P\Sigma P]. \qquad (3.11)$$

Proof. See Appendix A. $\square$

Remark 3.1. Note that it follows from Lemma 2.2 that if $\tilde{A}$ is asymptotically stable then conditions (3.2)-(3.5) are automatically satisfied.

Remark 3.2. Note that when P_2 and Q_2 are full rank, i.e., the controller (A_c, B_c, C_c) is minimal, the choices of $X, Y,$ and Z have no effect on (A_c, B_c, C_c).

Next, we specialize Theorem 3.1 to the case where (A_c, B_c, C_c) is a minimal reduced-order compensator.

Corollary 3.1. Suppose P_2 and Q_2 in Theorem 3.1 have full rank, i.e., rank $P_2 =$ rank $Q_2 = n_c \leq n$. Then, there exist nonnegative-definite matrices $P, Q, \hat{P}$ and $\hat{Q}$ such that A_c, B_c and C_c are given by

$$A_c = \Gamma(A - Q\bar{\Sigma} - \Sigma P)G^\mathrm{T}, \qquad (3.12a)$$

$$B_c = \Gamma Q C^\mathrm{T} V_2^{-1}, \qquad (3.12b)$$

$$C_c = R_2^{-1} B^\mathrm{T} P G^\mathrm{T}, \qquad (3.12c)$$

13

for some projective factorization (G, M, Γ) of $\hat{Q}\hat{P}$ and such that the following conditions are satisfied:

$$0 = A^{\mathrm{T}}P + PA + R_1 - P\Sigma P + \tau_\perp^{\mathrm{T}} P\Sigma P \tau_\perp, \tag{3.13}$$

$$0 = AQ + QA^{\mathrm{T}} + V_1 - Q\bar{\Sigma}Q + \tau_\perp Q\bar{\Sigma}Q\tau_\perp^{\mathrm{T}}, \tag{3.14}$$

$$0 = (A - Q\bar{\Sigma})^{\mathrm{T}}\hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P - \tau_\perp^{\mathrm{T}} P\Sigma P \tau_\perp, \tag{3.15}$$

$$0 = (A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^{\mathrm{T}} + Q\bar{\Sigma}Q - \tau_\perp Q\bar{\Sigma}Q\tau_\perp^{\mathrm{T}}, \tag{3.16}$$

$$\operatorname{rank}\hat{Q} = \operatorname{rank}\hat{P} = \operatorname{rank}\hat{Q}\hat{P} = n_c, \tag{3.17}$$

$$\tau = (\hat{Q}\hat{P})(\hat{Q}\hat{P})^{\#}. \tag{3.18}$$

Furthermore, the minimal cost is given by

$$J(A_c, B_c, C_c) = \operatorname{tr}[PV_1 + Q(P\Sigma P - \tau_\perp^{\mathrm{T}} P\Sigma P \tau_\perp)], \tag{3.19}$$

or, equivalently,

$$J(A_c, B_c, C_c) = \operatorname{tr}[QR_1 + P(Q\bar{\Sigma}Q - \tau_\perp Q\bar{\Sigma}Q\tau_\perp^{\mathrm{T}})]. \tag{3.20}$$

Proof. The proof is a direct consequence of Theorem 3.1. For details see [6]. □

Remark 3.3. Equations (3.13)-(3.18) are the standard optimal projection equations for reduced-order dynamic compensation given in [6].

Finally, we present a partial converse of the necessary conditions that guarantee closed-loop stability.

Corollary 3.2. Suppose there exist nonnegative-definite matrices $P, Q, \hat{P}$, and $\hat{Q}$ satisfying (3.13)-(3.18) and let A_c, B_c and C_c be given by (3.12). Then the compensator (A_c, B_c, C_c) is an extremal of the optimal fixed-order dynamic compensation problem. Furthermore the following are equivalent:

$$\tilde{A} \text{ is asymptotically stable;} \tag{3.21}$$

$$(\tilde{A}, \tilde{D}) \text{ is stabilizable;} \tag{3.22}$$

$$(\tilde{A}, \tilde{E}) \text{ is detectable.} \tag{3.23}$$

In addition,

$$(A_c, B_c) \text{ is controllable if and only if } A_c + B_c C G^{\mathrm{T}} \text{ is asymptotically stable,} \tag{3.24a}$$

14

(A_c, C_c) is observable if and only if $A_c + \Gamma B C_c$ is asymptotically stable. $\qquad$ (3.24b)

Proof. That the compensator (A_c, B_c, C_c) is an extremal follows immediately from the proof of Theorem 3.1. It follows from the proof of Corollary 3.1 that if nonegative-definite $P, Q, \hat{P}$ and $\hat{Q}$ satisfy (3.13)-(3.18) and the compensator (A_c, B_c, C_c) is given by (3.12), then (independent of the stability of $\tilde{A}$) there exist $\tilde{n} \times \tilde{n}$ real matrices $\tilde{P}$ and $\tilde{Q}$ satisfying (2.11), (2.12) and having partitioned forms (2.15), (2.16) with the partitions satisfying

$$P_1 = P + \hat{P}, \quad Q_1 = Q + \hat{Q}, \qquad (3.25a,b)$$

$$P_{12} = -\hat{P}G^{\mathrm{T}}, \quad Q_{12} = \hat{Q}\Gamma^{\mathrm{T}}, \qquad (3.25c,d)$$

$$P_2 = G\hat{P}G^{\mathrm{T}}, \quad Q_2 = \Gamma\hat{Q}\Gamma^{\mathrm{T}} \qquad (3.25e,f)$$

It then follows that $\tilde{P}$ and $\tilde{Q}$ can be expressed as

$$\tilde{P} = \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -I_n \\ G \end{bmatrix} \hat{P}[-I_n \quad G^{\mathrm{T}}], \qquad (3.26)$$

$$\tilde{Q} = \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} I_n \\ \Gamma \end{bmatrix} \hat{P}[I_n \quad \Gamma^{\mathrm{T}}], \qquad (3.27)$$

and thus

$$\tilde{P} \geq 0, \quad \tilde{Q} \geq 0. \qquad (3.28)$$

Obviously (3.21) implies (3.22) and (3.23). Conversely, using (2.11) and (2.12) it follows from Lemma 12.2 of [19] that (3.22) and (3.23) imply (3.20). Next, it follows from (3.25f) that the (2,2) block Q_2 of $\tilde{Q}$ satisfies

$$Q_2 = \Gamma\hat{Q}\Gamma^{\mathrm{T}} > 0. \qquad (3.29)$$

Furthermore, as shown in the proof of Theorem 3.1, Q_2 satisfies

$$0 = (A_c + B_c C G^{\mathrm{T}})Q_2 + Q_2(A_c + B_c C G^{\mathrm{T}})^{\mathrm{T}} + B_c V_2 B_c^{\mathrm{T}}. \qquad (3.30)$$

The equivalence (3.24a) then follows from (3.29), (3.30), Theorem 3.6 and Lemma 12.2 of [19]. The proof of (3.24b) follows in similar fashion by noting that $P_2 = G\hat{P}G^{\mathrm{T}} > 0$ and P_2 satisfies

$$0 = (A_c + \Gamma B C_c)^{\mathrm{T}} P_2 + P_2(A_c + \Gamma B C_c) + C_c^{\mathrm{T}} R_2 C_c. \quad \square$$

15

4. Optimality Conditions for Full-Order Dynamic Compensation

In this section, we restrict our attention to n^{th}-order compensators. Specifically, we show that even when the compensator is nonminimal, the generalized fixed-order equations of Theorem 3.1 always yield the standard LQG observer and regulator Riccati equations. We also show that a corresponding set of mutually coupled equations also exist that are identical in form to standard optimal projection equations but characterize the same compensator as obtained from the standard LQG Riccati equations. The proofs of these results rely on the balanced basis described in the following lemma.

Lemma 4.1. Consider the closed-loop system defined in (2.3). Suppose $n_c = n$ and $\hat{P}$ and $\hat{Q}$ are defined as in (2.36) and (2.37) with

$$\text{rank } \hat{P} = \text{rank } P_2 = n_p, \tag{4.1a}$$

$$\text{rank } \hat{Q} = \text{rank } Q_2 = n_q, \tag{4.1b}$$

and

$$\text{rank } \hat{Q}\hat{P} = \text{rank } Q_2 P_2 = n_r. \tag{4.1c}$$

If

$$P_{12}^T Q_{12} + P_2 Q_2 = 0, \tag{4.2}$$

then there exists a nonsingular $\tilde{n} \times \tilde{n}$ matrix

$$\tilde{S} = \begin{bmatrix} \dot{S}_1 & 0 \\ 0 & S_2 \end{bmatrix}, \quad S_1, S_2 \in \mathrm{IR}^{n \times n}, \tag{4.3}$$

such that

$$S_1^T \hat{P} S_1 = S_2^T P_2 S_2 = \begin{bmatrix} \Sigma_1 & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & \Sigma_2 & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{n_t \times n_r} & 0 & 0 & 0 \end{bmatrix}, \tag{4.4a}$$

$$S_1^{-1} \hat{Q} S_1^{-T} = S_2^{-1} Q_2 S_2^{-T} = \begin{bmatrix} \Sigma_1 & 0_{n_r \times (n_p - n_r)} & 0_{n_r \times (n_q - n_r)} & 0_{n_r \times n_t} \\ 0_{(n_p - n_r) \times n_r} & 0 & 0 & 0 \\ 0_{(n_q - n_r) \times n_r} & 0 & \Sigma_3 & 0 \\ 0_{n_t \times n_r} & 0 & 0 & 0 \end{bmatrix}, \tag{4.4b}$$

where $n_t = n - (n_p + n_q - n_r)$ and $\Sigma_1 \in \mathrm{IR}^{n_r \times n_r}$, $\Sigma_2 \in \mathrm{IR}^{(n_p - n_r) \times (n_p - n_r)}$, $\Sigma_3 \in \mathrm{IR}^{(n_q - n_r) \times (n_q - n_r)}$ are diagonal and positive definite.

16

Proof. It follows from (4.2) and statement (*ii*) of Lemma 2.5 that $\hat{Q}\hat{P}$ and $Q_2 P_2$ have the same non-zero eigenvalues. Thus using statement (*iv*) of Lemma 2.4 and the rank conditions specified in (4.1), yields (4.4a) and (4.4b). $\square$

Below, unless otherwise specified, all the $n \times n$ partitioned matrices have the same sub-matrix dimensions as in (4.4).

Definition 4.1. Suppose the closed-loop system (2.3) satisfies the conditions of Lemma 4.1 and is transformed via the similarity transformation $\tilde{S}$ given by (4.3) so that the new closed-loop states are given by $\tilde{x}'(t) = \tilde{S}^{-1}\tilde{x}(t)$, and hence the transformed plant triple (A', B', C') and compensator triple (A'_c, B'_c, C'_c) are given by $A' = S_1^{-1}AS_1$, $B' = S_1^{-1}B$, $C' = CS_1$, and $A'_c = S_2^{-1}A_c S_2$, $B'_c = S_2^{-1}B_c$, $C'_c = C_c S_2$. Furthermore, let the transformed closed-loop covariance $\tilde{Q}'$ and its dual $\tilde{P}'$ be given by $\tilde{Q}' = \tilde{S}^{-1}\tilde{Q}\tilde{S}^{-T}$, $\tilde{P}' = \tilde{S}^{T}\tilde{P}\tilde{S}$, so that

$$\hat{P}' = P_2' = \begin{bmatrix} \Sigma_1 & 0 & 0 & 0 \\ 0 & \Sigma_2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad \hat{Q}' = Q_2' = \begin{bmatrix} \Sigma_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \Sigma_3 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{4.5a, b}$$

Then the transformation $\tilde{S}$ is called a *strictly balanced transformation* and the transformed coordinates $\tilde{x}'$ are called *strictly balanced coordinates.*

Definition 4.2. Suppose the closed-loop system (2.3) is transformed via a similarity transformation $\hat{S} = \begin{bmatrix} S_1 & 0 \\ 0 & S_1 \end{bmatrix}$, where S_1 is as in (4.4) so that the new closed-loop states are given by $\tilde{x}'(t) = \hat{S}^{-1}\tilde{x}(t)$. In this case, the transformation $\hat{S}$ is called a *balanced transformation* and the transformed coordinates $\tilde{x}'$ are called *balanced coordinates.*

Theorem 4.1. Let $n_c = n$. Then there exist $n \times n$ nonnegative definite matrices $P_L, Q_L, \hat{P}_L$, and $\hat{Q}_L$ such that an admissible extremal of the full-order dynamic compensation problem is given by:

$$A_{CL} = A - \Sigma P_L - Q_L \bar{\Sigma}, \tag{4.6a}$$

$$B_{CL} = Q_L C^{T} V_2^{-1}, \tag{4.6b}$$

$$C_{CL} = R_2^{-1} B^{T} P_L, \tag{4.6c}$$

where P_L and Q_L are the unique, nonnegative definite solutions respectively of

$$0 = A^{T} P_L + P_L A + R_1 - P_L \Sigma P_L, \tag{4.7}$$

17

$$0 = AQ_L + Q_L A^T + V_1 - Q_L \bar{\Sigma} Q_L, \tag{4.8}$$

and $\hat{P}_L$ and $\hat{Q}_L$ satisfy

$$0 = (A - Q_L\bar{\Sigma})^T \hat{P}_L + \hat{P}_L(A - Q_L\bar{\Sigma}) + P_L\Sigma P_L, \tag{4.9}$$

$$0 = (A - \Sigma P_L)\hat{Q}_L + \hat{Q}_L(A - \Sigma P_L)^T + Q_L\bar{\Sigma}Q_L. \tag{4.10}$$

In this case, the minimal cost is given by

$$J(A_c, B_c, C_c) = \text{tr}[P_L V_1 + Q_L P_L \Sigma P_L], \tag{4.11}$$

or, equivalently,

$$J(A_c, B_c, C_c) = \text{tr}[Q_L R_1 + P_L Q_L \bar{\Sigma} Q_L]. \tag{4.12}$$

Proof. The proof is constructive in nature and follows from Theorem 3.1 by choosing $X = Q_L C^T V_2^{-1}$, $Y = R_2^{-1} B^T P_L$, and $Z = A - Q_L\bar{\Sigma} - \Sigma P_L$. For details see [14]. $\square$

Remark 4.1. Note that (4.7) and (4.8) are the standard decoupled regulator and observer Riccati equations and (A_{CL}, B_{CL}, C_{CL}) represents the LQG compensator (minimal or nonminimal) obtained through the fixed-structure approach. Furthermore, note that equations (4.9) and (4.10) are superfluous since the optimal compensator only depends upon the variables P_L and Q_L. However, using (4.6) it can be easily shown that $\hat{P}_L$ and $\hat{Q}_L$ are observability and controllability Gramians of the compensator [10].

Corollary 4.1. Suppose the compensator obtained in Theorem 4.1 is nonminimal. For convenience, let $n_q \triangleq \text{rank } \hat{Q}_L$, $n_p \triangleq \text{rank } \hat{P}_L$ and $n_r \triangleq \text{rank } \hat{Q}_L\hat{P}_L$. Then the compensator matrices in the balanced coordinates, $A'_{CL}, B'_{CL}, C'_{CL}$, have the following structure:

$$A'_{CL} = \begin{bmatrix} A'_{CL,11} & A'_{CL,12} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ A'_{CL,31} & A'_{CL,32} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{4.13a}$$

$$B'_{CL} = \begin{bmatrix} B'_{CL,1} \\ 0 \\ B'_{CL,3} \\ 0 \end{bmatrix}, \tag{4.13b}$$

$$C'_{CL} = [C'_{CL,1} \quad C'_{CL,2} \quad 0 \quad 0], \tag{4.13c}$$

18

where

$$A'_{CL,11} \in \mathrm{IR}^{n_r \times n_r}, A'_{CL,12} \in \mathrm{IR}^{n_r \times (n_p - n_r)}, A'_{CL,31} \in \mathrm{IR}^{(n_q - n_r) \times n_r}, A'_{CL,32} \in \mathrm{IR}^{(n_q - n_r) \times (n_p - n_r)},$$

$$B'_{CL,1} \in \mathrm{IR}^{n_r \times l}, B'_{CL,3} \in \mathrm{IR}^{(n_q - n_r) \times l},$$

$$C'_{CL,1} \in \mathrm{IR}^{m \times n_r}, C'_{CL,2} \in \mathrm{IR}^{m \times (n_p - n_r)}.$$

Proof. The proof is a direct consequence of Theorem 4.1 and relies on transforming the compensator (4.6) into strictly balanced coordinates. □

Next, using the balanced transformation presented in Definition 4.1, we show that the input-output map of the nonminimal LQG compensator given by (4.6) or, equivalently, (4.13) is equivalent to the input-output map of a specific compensator of Theorem 3.1 with $X = 0$, $Y = 0$, $Z = 0$ and $n_c = n$, which we shall call the *full-order central compensator* or the *full-order least-squares compensator*. In this case, as shown in the next theorem, the resulting optimality conditions are identical in structure to the standard optimal projection equations given by Corollary 3.1.

Theorem 4.2. Let $n_c = n$ and let n_l represent the order of the minimal realization of the LQG controller. Suppose rank $Q_2 =$ rank $P_2 =$ rank $Q_2 P_2 = n_r$, where $Q_2 \in \mathrm{IR}^{n \times n}$ and $P_2 \in \mathrm{IR}^{n \times n}$ are respectively the closed-loop covariance of the controller states and its dual. Then there exist $n \times n$ nonnegative-definite matrices $P, Q, \hat{P}$, and $\hat{Q}$ such that an extremal of the full-order compensation problem is given by

$$A_c = \Gamma_F (A - Q\bar{\Sigma} - \Sigma P) G_F^\mathrm{T}, \tag{4.14a}$$

$$B_c = \Gamma_F Q C^\mathrm{T} V_2^{-1}, \tag{4.14b}$$

$$C_c = R_2^{-1} B^\mathrm{T} P G_F^\mathrm{T}, \tag{4.14c}$$

where $P, Q, \hat{P}$, and $\hat{Q}$ satisfy:

$$0 = A^\mathrm{T} P + P A + R_1 - P\Sigma P + \tau_\perp^\mathrm{T} P\Sigma P\tau_\perp, \tag{4.15}$$

$$0 = AQ + QA^\mathrm{T} + V_1 - Q\bar{\Sigma}Q + \tau_\perp Q\bar{\Sigma}Q\tau_\perp^\mathrm{T}, \tag{4.16}$$

$$0 = (A - Q\bar{\Sigma})^\mathrm{T}\hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P - \tau_\perp^\mathrm{T} P\Sigma P\tau_\perp, \tag{4.17}$$

$$0 = (A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^\mathrm{T} + Q\bar{\Sigma}Q - \tau_\perp Q\bar{\Sigma}Q\tau_\perp^\mathrm{T}, \tag{4.18}$$

$$\text{rank } \hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_r, \tag{4.19}$$

19

$$\tau = (\hat{Q}\hat{P})(\hat{Q}\hat{P})^{\#}. \tag{4.20}$$

and (G_F, M_F, Γ_F) is a nth-order generalized projective factorization of $\hat{Q}\hat{P}$. In addition, the corresponding cost is given by

$$J(A_c, B_c, C_c) = \text{tr}[PV_1 + Q(P\Sigma P - \tau_{\perp}^{\mathrm{T}} P\Sigma P\tau_{\perp})], \tag{4.21}$$

or, equivalently,

$$J(A_c, B_c, C_c) = \text{tr}[QR_1 + P(Q\bar{\Sigma}Q - \tau_{\perp}Q\bar{\Sigma}Q\tau_{\perp}^{\mathrm{T}})]. \tag{4.22}$$

Furthermore, suppose that the compensator is nonminimal, i.e., $n_l < n$, the minimal dimension of the full-order central compensator is the same as that of the LQG compensator, i.e., $n_l = n_r$, and that the LQG compensator matrices for this plant in the balanced coordinates are as in (4.13) of Corollary 4.1. Then, in the balanced coordinates, a central controller triple (A_c', B_c', C_c'') is given by:

$$A_c' = \begin{bmatrix} A_{CL,11}' & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{4.23a}$$

$$B_c' = \begin{bmatrix} B_{CL,1}' \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{4.23b}$$

$$C_c' = [C_{CL,1}' \quad 0 \quad 0 \quad 0], \tag{4.23c}$$

where $A_{CL,11}', B_{CL,1}'$ and $C_{CL,1}'$ are as in Corollary 4.1.

Proof. The proof is similar to the proof of Theorem 3.1 with $n_c = n$ and $X = 0, Y = 0, Z = 0$. Additionally, the compensator (4.23) is a direct consequence of a balanced transformation. $\square$

Remark 4.2. Suppose the full-order compensator is minimal which implies rank $\hat{Q}$ = rank $\hat{P}$ = rank $\hat{Q}\hat{P} = n$, or, equivalently, $\tau = (\hat{Q}\hat{P})(\hat{Q}\hat{P})^{\#} = I_n$. Then, (4.15)-(4.18) reduce to (4.7)-(4.10). Furthermore, it follows from the identity $\Gamma_F = G_F^{-\mathrm{T}}$ that (A_c, B_c, C_c) of (4.14) is simply some similarity transformation of the LQG controller triple (A_{CL}, B_{CL}, C_{CL}) in (4.6).

Remark 4.3. Suppose the LQG compensator is nonminimal. It can easily be shown using the balanced realizations that the LQG controller (4.13) and the full-order central controller with the same minimal dimension (4.23) have the same input-output map (but are not equivalent within a change of basis). Furthermore, they have the same input-output map as an extremal of the optimal reduced-order compensator problem with $n_c = n_l$.

20

Next, we prove that the optimal reduced-order controller is also a projection of a full-order central controller whose minimal dimension is the same as the dimension of the optimal reduced-order controller.

Theorem 4.3. Suppose $n_c \leq n_l$ in Theorem 4.2. Then a n_c^{th}-order optimal compensator given by Corollary 3.1 can be attained through a projection $\mu \triangleq Q_{2F} P_{2F} (Q_{2F} P_{2F})^\#$ of a full-order central compensator with minimal dimension n_c, where Q_{2F} and P_{2F} are respectively the closed-loop covariance of the full-order central controller states and its dual.

Proof. Let $n_r = n_c$ in Theorem 4.2 and note that (4.15)-(4.20) are identical to (3.13)-(3.18) which implies that $P, Q, \hat{P}, \hat{Q}$ and τ are identical for a full-order central compensator and a n_c^{th}-order compensator obtained from Corollary 3.1.

Next, let Q_{2F} and P_{2F} be respectively the closed-loop covariance of the full-order central controller states and its dual. Since $Q_{2F} \geq 0, P_{2F} \geq 0$, and rank $Q_{2F} =$ rank $P_{2F} =$ rank $Q_{2F} P_{2F} = n_c$, it follows from Definition 2.3 and [17] that there exists a projective factorization (T_r^T, M_r, L_r) with $T_r \in \mathbb{R}^{n \times n_c}, L_r \in \mathbb{R}^{n_c \times n}$ and $M_r \in \mathbb{R}^{n_c \times n_c}$ such that

$$Q_{2F} P_{2F} = T_r M_r L_r, \tag{4.24}$$

$$L_r T_r = I_{n_c}, \tag{4.25}$$

$$\mu \triangleq Q_{2F} P_{2F} (Q_{2F} P_{2F})^\# = T_r L_r, \tag{4.26}$$

and μ is a $n \times n$ projection matrix.

Now it follows from (A.4) and the identities $G_F \triangleq Q_2^\dagger Q_{12}^T$, $M_F \triangleq Q_2 P_2$, and $\Gamma_F \triangleq -P_2^\dagger P_{12}^T$, that

$$L_r \Gamma_F G_F^T T_r = L_r P_{2F}^\dagger P_{2F} Q_{2F} Q_{2F}^\dagger T_r. \tag{4.27}$$

Next, noting the property of the projective factorization that $\mathcal{R}(L_r^T) = \mathcal{R}(P_{2F})$ and $\mathcal{R}(T_r) = \mathcal{R}(Q_{2F})$, it follows from property *(ii)* of Lemma 2.1 that $L_r P_{2F}^\dagger P_{2F} = L_r$ and $Q_{2F} Q_{2F}^\dagger T_r = T_r$. Using (4.25), (4.27) can be rewritten as

$$L_r \Gamma_F G_F^T T_r = L_r T_r = I_{n_c}. \tag{4.28}$$

In addition, using (4.26) and $Q_{2F} P_{2F} (Q_{2F} P_{2F})^\# Q_{2F} P_{2F} = Q_{2F} P_{2F}$, yields

$$T_r L_r M_F T_r L_r = Q_{2F} P_{2F} = M_F,$$

21

which implies

$$G_F^{\mathrm{T}} T_r L_r M_F T_r L_r \Gamma_F = G_F^{\mathrm{T}} M_F \Gamma_F = \hat{Q}\hat{P}. \tag{4.29}$$

Thus, it follows from (4.28) and (4.29) that $(T_r^{\mathrm{T}} G_F, L_r M_F T_r, L_r \Gamma_F)$ is a projective factorization of $\hat{Q}\hat{P}$. Hence, using Corollary 3.1, (A_r, B_r, C_r) with

$$A_r = L_r \Gamma_F (A - Q\bar{\Sigma} - \Sigma P) G_F^{\mathrm{T}} T_r, \tag{4.30a}$$

$$B_r = L_r \Gamma_F Q C^{\mathrm{T}} V_2^{-1}, \tag{4.30b}$$

$$C_r = R_2^{-1} B^{\mathrm{T}} P G_F^{\mathrm{T}} T_r. \tag{4.30c}$$

is a n_c^{th}-order optimal compensator. However, noting (4.14), (A_r, B_r, C_r) can also be expressed as $(L_r A_c T_r, L_r B_c, C_c T_r)$. Thus, it follows from (4.23) and (4.24) that a n_c^{th}-order optimal compensator (A_r, B_r, C_r) can be obtained through the projection μ of a full-order central compensator (A_c, B_c, C_c). $\square$

The balanced controller reduction method of [10] characterizes the reduced-order controller by a projection of the LQG controller. For the special case in which the LQG controller is nonminimal and the requested dimension of the reduced-order controller equals the minimal dimension of the LQG controller, i.e., $n_c = n_l$, this method is capable of producing a minimal representation of the LQG compensator. For this special case, the following theorem explicitly defines the relationships among the projection matrix ν used by the suboptimal balanced controller reduction method, the projection matrix μ given by Theorem 4.3 through which a central compensator of rank n_l is projected into a n_l^{th}-order optimal compensator and the optimal projection matrix τ from standard optimal projection theory. To facilitate the exposition of the following theorem, following Definitions 4.2 and 4.1, let $\begin{bmatrix} S_{1,L} & 0 \\ 0 & S_{1,L} \end{bmatrix}$ denote the balanced transformation of the closed-loop system using the LQG compensator and let $\begin{bmatrix} S_{1,c} & 0 \\ 0 & S_{1,c} \end{bmatrix}$ and $\begin{bmatrix} S_{1,c} & 0 \\ 0 & S_{2,c} \end{bmatrix}$ denote respectively the balanced transformation and strictly balanced transformation of the closed-loop system using the appropriate central compensator of rank n_l.

Theorem 4.4. Suppose (A_{CL}, B_{CL}, C_{CL}) is a nonminimal LQG compensator with minimal dimension n_l and (A_c, B_c, C_c) is an appropriate central compensator with rank n_l. Let ν and μ be $n \times n$ projection matrices of rank n_l and $L_\nu, L_\mu \in \mathbb{R}^{n_l \times n}$ and $T_\nu, T_\mu \in \mathbb{R}^{n \times n_l}$ satisfy $\nu = T_\nu L_\nu$, $\mu = T_\mu L_\mu$ and $L_\nu T_\nu = L_\mu T_\mu = I_{n_l}$. Suppose that $(L_\nu A_{CL} T_\nu, L_\nu B_{CL}, C_{CL} T_\nu)$ and $(L_\mu A_c T_\mu, L_\mu B_c, C_c T_\mu)$ are minimal realizations of (A_{CL}, B_{CL}, C_{CL}). Then

$$\nu = T^{-1} \mu T. \tag{4.31}$$

22

where $T = S_{1,L}^{-1} S_{2,c}$. In addition,

$$\nu = W^{-1} \tau W, \tag{4.32}$$

where $W = S_{1,L}^{-1} S_{1,c}$.

Proof. The LQG compensator triple in balanced coordinates, $(A'_{CL} = S_{1,L}^{-1} A_{CL} S_{1,L}, \ B'_{CL} = S_{1,L}^{-1} B_{CL}, \ C'_{CL} = C_{CL} S_{1,L})$ has the expression as in (4.13). Using (4.13), the minimal representation of the LQG controller in balanced coordinates is $(A'_{CL,11}, B'_{CL,1}, C'_{CL,1})$ which implies that $\nu' = \begin{bmatrix} I_{n_l} & 0 \\ 0 & 0 \end{bmatrix}$ in balanced coordinates and $L'_\nu = [I_{n_l} \ 0]$, and $T'_\nu = [I_{n_l} \ 0]^{\mathrm{T}}$ is a factorization of ν'. Thus, using the following identities

$$A'_{CL,11} = L'_\nu A'_{CL} T'_\nu = L'_\nu S_{1,L}^{-1} A_{CL} S_{1,L} T'_\nu = L_\nu A_{CL} T_\nu,$$

$$B'_{CL,1} = L'_\nu B'_{CL} = L'_\nu S_{1,L}^{-1} B_{CL}, \ \text{and} \ C'_{CL,1} = C'_{CL} T'_\nu = C_{CL} S_{1,L} T'_\nu,$$

yields

$$L_\nu = L'_\nu S_{1,L}^{-1}, \ T_\nu = S_{1,L} T'_\nu,$$

and

$$\nu = S_{1,L} \nu' S_{1,L}^{-1} = S_{1,L} \begin{bmatrix} I_{n_l} & 0 \\ 0 & 0 \end{bmatrix} S_{1,L}^{-1}. \tag{4.33}$$

Next, noting (4.4) and using properties *(iii)* and *(iv)* of Lemma 2.4, we obtain

$$\nu = (\hat{Q}_L \hat{P}_L)(\hat{Q}_L \hat{P}_L)^{\#}. \tag{4.34}$$

Now, according to Remark 4.3, a n_l^{th}-order optimal controller obtained from (3.12)-(3.18) has the same input-output map as the LQG controller whose minimal dimension is n_l. Thus, using the special case, $n_c = n_l$, of Theorem 4.3, yields

$$\mu = (Q_{2F} P_{2F})(Q_{2F} P_{2F})^{\#}, \tag{4.35}$$

where Q_{2F} and P_{2F} are respectively the closed-loop covariance of the appropriate n_l-rank central controller states and its dual. Now it follows from Definition 4.1 that

$$\mu = S_{2,c} \begin{bmatrix} I_{n_l} & 0 \\ 0 & 0 \end{bmatrix} S_{2,c}^{-1}. \tag{4.36}$$

Finally, following (4.20) and Definition 4.2, yields

$$\tau = \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#} = S_{1,c} \begin{bmatrix} I_{n_l} & 0 \\ 0 & 0 \end{bmatrix} S_{1,c}^{-1}. \tag{4.37}$$

Using (4.33), (4.36) and (4.37), yields (4.31) and (4.32). $\quad\square$

23

5. Numerical Solution of the Coupled Design Equations and Illustrative Results

One of the principal motivations for the Riccati equation approach to reduced-order dynamic compensation is the opportunity it provides for developing efficient computational algorithms for control design. In particular, the goal has been to develop numerical methods which exploit the structure of the Riccati equations. It turns out however, that methods for solving standard Riccati equations cannot account for the additional terms appearing in the modified equations of Theorem 3.1 and Corollary 3.1. Therefore, a new class of numerical algorithms has recently been developed based upon homotopic continuation methods. These methods operate by first replacing the original problem by a simpler problem with a known solution. Specifically, the simpler problem can be chosen to correspond to a low authority full-order LQG control problem. As shown in [20], if the weighting matrices are chosen properly, then in this case the LQG compensator is nearly nonminimal. Hence, using a simple balanced controller reduction technique [10], the resulting balanced reduced-order controller serves as a good approximation to the optimal projection controller corresponding to the simpler problem. The desired solution is then reached by integrating along a path which connects the starting problem to the original problem. These ideas have been recently illustrated for the reduced-order control problem in [13, 14].

Using the homotopy algorithm appearing in [13] we demonstrate the utility of the H_2 optimal reduced-order controller design framework discussed in this paper on the four-disk axial beam problem shown in Figure 5.1. This example was derived from a laboratory experiment [21] and has been considered in several subsequent publications (e.g., [22-24]). The 8th-order state space model, problem data, and design weights are given in the above references. The basic control objective for the four-disk problem is to control the angular displacement at the location of disk 1 using a torque input at the location of disk 3. It is also assumed that a torque disturbance enters the system at the location of disk 3.

The design philosophy adopted here is that the scaling q_2 of the nominal control weight $R_{2,0} = 1$ and the nominal sensor noise intensity $V_{2,0} = 1$, where the subscript "0" denotes initial values, are simply "design knobs" used to determine the control authority. Hence, $R_2(\lambda) = q_2(\lambda)R_{2,0}$ and $V_2(\lambda) = q_2(\lambda)V_{2,0}$, where λ is a homotopy parameter and $\lambda \in [0,1]$. Here, we consider the design of $2^{nd}, 4^{th}$, and 6th -order controllers for various authority levels.

Since at $q_2 = 10$, the $2^{nd}, 4^{th}$, and 6th balanced reduced-order controllers are all good approximations of the corresponding reduced-order optimal controllers, we use these suboptimal controllers

24

to initialize the homotopy algorithm and deform the controllers into the higher authority optimal controllers corresponding to $q_2 = 1$. In each of the following passes, we increase the authority level by decreasing R_2 and V_2 by a factor of 10, i.e., $q_{2,\text{next}} = 0.1q_{2,0}$, and at the end of each pass deform the initial optimal controllers to the optimal controllers corresponding to the higher authority level. This process is repeated for every reduced-order design. Figure 5.2 compares the optimal controllers of various orders. This type of figure can be used in practice to determine the order of the controller to be implemented.

The Frobenius norms of $P, Q, \hat{P}$, and $\hat{Q}$ are also recorded along the homotopy path and typical results are shown in Figure 5.3 of $\|\hat{P}\|_F$ for the 4th-order controller design. It is interesting to note that as the control authority is increased beyond a certain level (e.g., for $n_c = 4, q_2 < 10^{-4}$) those values approach some stable limit as indicated in this figure. This is because $P, Q, \hat{P}$, and $\hat{Q}$ converge to fixed values as the control authority increases. It follows that the optimal reduced-order controller converges to a fixed value.

6. Conclusion

Necessary conditions for fixed-structure H_2 optimal control were derived without assuming compensator minimality. These necessary conditions are characterized in terms of coupled Riccati and Lyapunov equations which reduce to the optimal projection equations [6] when the compensator is minimal. The standard LQG Riccati equations can also be derived when the optimality conditions are specialized to the full-order case. Furthermore, it is shown for the first time that a reduced-order optimal projection controller is a projection of a "central" extremal of the corresponding full-order compensation problem. For nonminimal LQG compensators, balanced controller reduction method is able to produce a minimal-order realization of the LQG compensator. For this special case, the relationships between the projection matrix used by balanced controller reduction, the projection matrix through which an appropriate central controller is projected into a reduced-order optimal controller whose dimension is the same as the minimal dimension of the LQG controller, and the optimal projection matrix from the standard optimal projection theory are explicitly defined. A continuation algorithm that exploits the Riccati-equation design framework is discussed and its utility for controller synthesis is illustrated using a representative problem in structural control.

25

Appendix A. *Proof of Theorem 3.1*

To optimize (2.14) over the open set $\mathcal{S}_c$ subject to the constraint (2.11), form the Lagrangian

$$\mathcal{L}(A_c, B_c, C_c, \tilde{Q}, \tilde{P}, \lambda) \triangleq \text{tr}[\lambda \tilde{Q} \tilde{R} + (\tilde{A} \tilde{Q} + \tilde{Q} \tilde{A}^{\text{T}} + \tilde{V}) \tilde{P}], \qquad (A.1)$$

where the Lagrange multipliers $\lambda \geq 0$ and $\tilde{P} \in \text{IR}^{\tilde{n} \times \tilde{n}}$ are not both zero. We thus obtain

$$\frac{\partial \mathcal{L}}{\partial \tilde{Q}} = \tilde{A}^{\text{T}} \tilde{P} + \tilde{P} \tilde{A} + \lambda \tilde{R}.$$

Setting $\partial \mathcal{L} / \partial \tilde{Q} = 0$ yields

$$0 = \tilde{A}^{\text{T}} \tilde{P} + \tilde{P} \tilde{A} + \lambda \tilde{R},$$

or, equivalently,

$$(\tilde{A}^{\text{T}} \oplus \tilde{A}^{\text{T}}) \text{ vec } \tilde{P} = -\lambda \text{ vec } \tilde{R}, \qquad (A.2)$$

where $\oplus$ denotes the Kronecker sum and "vec" is the column-stacking operation defined in [25]. Since $\tilde{A}$ is assumed to be asymptotically stable, $(\tilde{A}^{\text{T}} \oplus \tilde{A}^{\text{T}})$ is invertible, and thus $\lambda = 0$ implies $\tilde{P} = 0$. Hence, it can be assumed without loss of generality that $\lambda = 1$, which yields (2.12). Furthermore, with $\lambda = 1$, (A.2) is equivalent to

$$\tilde{P} = -\text{vec}^{-1}[(\tilde{A}^{\text{T}} \oplus \tilde{A}^{\text{T}})^{-1} \text{vec } \tilde{R}].$$

To prove that $\tilde{P}$ is nonnegative definite, we rewrite the above expression as

$$\tilde{P} = \int_0^\infty \text{vec}^{-1}[e^{(\tilde{A}^{\text{T}} \oplus \tilde{A}^{\text{T}})t} \text{vec } \tilde{R}] \, dt, \qquad (A.3)$$

and show that the integrand is nonnegative definite for all $t \in [0, \infty)$. For convenience, let S and N be $n \times n$ matrices with $N \geq 0$. Since (see [25])

$$e^{S \oplus S} = e^S \otimes e^S,$$

and

$$\text{vec}^{-1}[(S \otimes S)\text{vec } N] = SNS^{\text{T}} \geq 0,$$

where $\otimes$ denotes the Kronecker product defined in [25], it follows that

$$\text{vec}^{-1}[e^{S \oplus S} \text{vec } N] = \text{vec}^{-1}[(e^S \otimes e^S)\text{vec } N] = e^S N e^{S^{\text{T}}} \geq 0.$$

26

Noting $\tilde{R}$ is nonnegative definite and applying the above expression with (A.3) it follows that the integrand of (A.3) is nonnegative definite. Thus $\tilde{P}$ is nonnegative definite.

Now partition $\tilde{n} \times \tilde{n}$ $\tilde{P}, \tilde{Q}$ into $n \times n, n \times n_c$, and $n_c \times n_c$ subblocks as in (2.16) and (2.15), respectively. The stationary conditions, with $\lambda = 1$, are then given by

$$\frac{\partial \mathcal{L}}{\partial A_c} = P_{12}^{\mathrm{T}} Q_{12} + P_2 Q_2 = 0, \qquad (A.4)$$

$$\frac{\partial \mathcal{L}}{\partial B_c} = P_2 B_c V_2 + (P_{12}^{\mathrm{T}} Q_1 + P_2 Q_{12}^{\mathrm{T}}) C^{\mathrm{T}} = 0, \qquad (A.5)$$

$$\frac{\partial \mathcal{L}}{\partial C_c} = -R_2 C_c Q_2 + B^{\mathrm{T}} (P_1 Q_{12} + P_{12} Q_2) = 0. \qquad (A.6)$$

Expanding (2.11) and (2.12) yields (2.17)-(2.22). Since $\tilde{A}$ is assumed to be asymptotically stable, using (2.28b) and (2.29a) of Lemma 2.3 we can rewrite (A.4) as

$$P_2(I_{n_c} + P_2^\dagger P_{12}^{\mathrm{T}} Q_{12} Q_2^\dagger) Q_2 = 0. \qquad (A.7)$$

Next, define the $n \times n$ matrices

$$P \triangleq P_1 - P_{12} P_2^\dagger P_{12}^{\mathrm{T}}, \quad Q \triangleq Q_1 - Q_{12} Q_2^\dagger Q_{12}^{\mathrm{T}}. \qquad (A.8a, b)$$

$$\hat{P} \triangleq P_{12} P_2^\dagger P_{12}^{\mathrm{T}}, \quad \hat{Q} \triangleq Q_{12} Q_2^\dagger Q_{12}^{\mathrm{T}}, \qquad (A.9a, b)$$

and $n_c \times n, n_c \times n_c$, and $n_c \times n$ matrices

$$G \triangleq Q_2^\dagger Q_{12}^{\mathrm{T}}, \quad M \triangleq Q_2 P_2, \quad \Gamma \triangleq -P_2^\dagger P_{12}^{\mathrm{T}}. \qquad (A.10a, b, c)$$

Note that the definitions of $\hat{P}, \hat{Q}, G, M$ and Γ in (A.9) and (A.10) are identical to the ones defined in (2.41)-(2.45) of Lemma 2.5. Furthermore, noting that (A.4) is equivalent to (2.46), it follows from the property (ii) of Lemma 2.5 that (G, M, Γ) satisfies property (v) of Lemma 2.4, or, equivalently, (G, M, Γ) is a generalized projective factorization of $\hat{Q}\hat{P}$. Clearly, $P, Q, \hat{P}$ and $\hat{Q}$ are symmetric and $\hat{P}$ and $\hat{Q}$ are nonnegative-definite. To show that P and Q are also nonnegative definite, note that P is the upper left-hand block of the nonnegative-definite matrix $\tilde{\mathcal{P}} \tilde{P} \tilde{\mathcal{P}}^{\mathrm{T}}$, where

$$\tilde{\mathcal{P}} \triangleq \begin{bmatrix} I_n & -P_{12} P_2^\dagger \\ 0_{n_c \times n} & I_{n_c} \end{bmatrix}.$$

Similarly, Q is nonnegative definite. Next, using the properties of the Moore-Penrose generalized inverse, it is helpful to note the following identities from the definitions (A.8)-(A.10):

$$\hat{P} = -P_{12}\Gamma = -\Gamma^{\mathrm{T}} P_{12}^{\mathrm{T}} = \Gamma^{\mathrm{T}} P_2 \Gamma, \qquad (A.11a)$$

27

$$\hat{Q} = Q_{12}G = G^{T}Q_{12}^{T} = G^{T}Q_2G, \qquad (A.11b)$$

and

$$P = P_1 - \hat{P}, \ \ Q = Q_1 - \hat{Q}. \qquad (A.12a, b)$$

In addition, using (2.23) and (2.24) yields

$$P_{12}^{T} = -P_2\Gamma, \ \ P_{12} = -\Gamma^{T}P_2. \qquad (A.13)$$

$$Q_{12}^{T} = Q_2G, \ \ Q_{12} = G^{T}Q_2. \qquad (A.14)$$

Furthermore, it follows from (A.2) that

$$\hat{Q}\hat{P} = -Q_{12}P_{12}^{T}. \qquad (A.15)$$

Next, using (A.13), we can rewrite (A.7) as

$$P_2(I_{n_c} - \Gamma G^{T})Q_2 = 0. \qquad (A.16)$$

Forming $P_{12}P_2^{\dagger}(A.4)$ and using (A.9a) and (2.28a) of Lemma 2.3, yields

$$\hat{P}Q_{12} + P_{12}Q_2 = 0. \qquad (A.17)$$

Similarly, computing $(A.4)Q_2^{\dagger}Q_{12}^{T}$ and using (A.9b) and (2.29b) of Lemma 2.3, yields

$$P_{12}^{T}\hat{Q} + P_2Q_{12}^{T} = 0. \qquad (A.18)$$

Using the identities (A.18) and (A.12a) we can rewrite (A.5) as $P_2B_cV_2 + P_{12}^{T}QC^{T} = 0$. Noting that V_2 is invertible and using (A.13) yields $P_2(B_c - \Gamma QC^{T}V_2^{-1}) = 0$, which further implies

$$B_c = \Gamma QC^{T}V_2^{-1} + (I_{n_c} - P_2^{\dagger}P_2)X, \qquad (A.19)$$

where $X \in \mathbb{R}^{n_c \times l}$ is an arbitrary matrix. Similarly, using (A.17) and (A.12b) we can rewrite (A.6) as $-R_2C_cQ_2 + B^{T}PQ_{12} = 0$. Noting that R_2 is invertible and using (A.14), yields $(C_c - R_2^{-1}B^{T}PG^{T})Q_2 = 0$, which further implies

$$C_c = R_2^{-1}B^{T}PG^{T} + Y(I_{n_c} - Q_2Q_2^{\dagger}), \qquad (A.20)$$

where $Y \in \mathbb{R}^{m \times n_c}$ is an arbitrary matrix. Next, computing either $(2.21)^{T}Q_{12} + (2.22)Q_2$ and using (A.4) and (A.6) or $P_{12}^{T}(2.18) + P_2(2.19)$ and using (A.4) and (A.5), yields

$$P_2A_cQ_2 + P_{12}^{T}AQ_{12} + P_2B_cCQ_{12} - P_{12}^{T}BC_cQ_2 = 0.$$

Noting (A.13) and (A.14), the above equation is equivalent to

$$P_2(A_c - \Gamma A G^{\mathrm{T}} + B_c C G^{\mathrm{T}} + \Gamma B C_c)Q_2 = 0. \tag{A.21}$$

Using identities (A.19) and (A.20) and noting the properties of the Moore-Penrose generalized inverse, (2.23) and (2.24), (A.21) can be rewritten as:

$$P_2[A_c - \Gamma(A - \Sigma P - Q\bar{\Sigma})G^{\mathrm{T}}]Q_2 = 0, \tag{A.22}$$

which further implies

$$A_c = \Gamma(A - \Sigma P - Q\bar{\Sigma})G^{\mathrm{T}} + Z - P_2^{\dagger} P_2 Z Q_2 Q_2^{\dagger}, \tag{A.23}$$

where $Z \in \mathrm{I\!R}^{n_c \times n_c}$ is arbitrary.

However, it follows from properties (ii), (iii), (v) and (vi) of Lemma 2.2, X, Y and Z must be chosen such that the compensator triple (A_c, B_c, C_c) satisfies (3.4), (3.5), $\mathcal{N}(P_2)$ = the unobservable subspace of (A_c, C_c), and $\mathcal{N}(Q_2)$ = the uncontrollable subspace of (A_c, B_c), to assure the closed-loop system stability.

Furthermore, since P_2 and Q_2 do not necessarily have full rank and $\mathcal{N}(P_2) \subset \mathcal{N}(P_{12})$ and $\mathcal{N}(Q_2) \subset \mathcal{N}(Q_{12})$, we may introduce additional singularities during the derivation for the expression of A_c. The extra singularities can be eliminated by examining the conditions under which the original Lyapunov equations (2.17)-(2.22) are satisfied. Since A_c is not involved in (2.17) and (2.20), only (2.18), (2.19), (2.21) and (2.22) have to be checked. Using (2.28) and (2.30), (2.21) can be rewritten as

$$P_{12} A_c + (A^{\mathrm{T}} P_{12} + C^{\mathrm{T}} B_c^{\mathrm{T}} P_2 - P_1 B C_c)P_2^{\dagger} P_2 = 0.$$

Using (2.21) the above equation can be reduced to $P_{12} A_c - P_{12} A_c P_2^{\dagger} P_2 = 0$, or, equivalently,

$$P_{12} A_c(I - P_2^{\dagger} P_2) = 0. \tag{A.24}$$

Similarly, using (2.28) and (2.30), (2.22) can be reduced to (3.2). Next, using (2.29) and (2.31), (2.18) and (2.19) can be reduced to

$$Q_{12} A_c^{\mathrm{T}}(I - Q_2^{\dagger} Q_2) = 0, \tag{A.25}$$

and (3.3), respectively. Using the property $\mathcal{N}(P_2) \subset \mathcal{N}(P_{12})$, (3.2) implies (A.24). Similarly, with $\mathcal{N}(Q_2) \subset \mathcal{N}(Q_{12})$, (3.3) implies (A.25). Furthermore, note that (3.2) and (3.3) also satisfy the

necessary conditions for $\tilde{A}$ to be asymptotically stable as stated in Lemma 2.2. Thus X, Y and Z in (3.1) must be chosen such that A_c, B_c and C_c satisfy (3.2)-(3.5).

Computing $(2.20) + (2.21)\Gamma + \Gamma^{\mathrm{T}}(2.21)^{\mathrm{T}}$ and using (A.11a), (A.12a) and (A.13), yields

$$0 = A^{\mathrm{T}}P + PA + R_1 - PBC_c\Gamma - \Gamma^{\mathrm{T}}C_c^{\mathrm{T}}B^{\mathrm{T}}P - \Gamma^{\mathrm{T}}(A_c^{\mathrm{T}}P_2 + P_2 A_c - P_{12}^{\mathrm{T}}BC_c - C_c^{\mathrm{T}}B^{\mathrm{T}}P_{12})\Gamma. \quad (A.26)$$

Using (2.22) to eliminate the terms in the parenthesis, (A.26) is equivalent to

$$0 = A^{\mathrm{T}}P + PA + R_1 - PBC_c\Gamma - \Gamma^{\mathrm{T}}C_c^{\mathrm{T}}B^{\mathrm{T}}P + \Gamma^{\mathrm{T}}C_c^{\mathrm{T}}R_2 C_c \Gamma. \quad (A.27)$$

Forming $(A.27) + P\Sigma P - P\Sigma P$, and noting

$$(R_2^{-1}B^{\mathrm{T}}P - C_c\Gamma)^{\mathrm{T}}R_2(R_2^{-1}B^{\mathrm{T}}P - C_c\Gamma) = -PBC_c\Gamma - \Gamma^{\mathrm{T}}C_c^{\mathrm{T}}B^{\mathrm{T}}P + \Gamma^{\mathrm{T}}C_c^{\mathrm{T}}R_2 C_c\Gamma + P\Sigma P,$$

yields (3.6). Next, computing $(2.20) - (3.6)$, using (A.12a) and noting the identity $P_{12}B_c C = -\Gamma^{\mathrm{T}}P_2 B_c C = -\hat{P}Q\bar{\Sigma}$, we obtain (3.8). Similarly, computing $(2.17) - (2.18)G - G^{\mathrm{T}}(2.18)^{\mathrm{T}}$ and using (A.11b), (A.12b) and (A.14) yields

$$0 = AQ + QA^{\mathrm{T}} + V_1 - QC^{\mathrm{T}}B_c^{\mathrm{T}}G - G^{\mathrm{T}}B_c CQ - G^{\mathrm{T}}(A_c Q_2 + Q_2 A_c^{\mathrm{T}} + B_c CQ_{12} + Q_{12}^{\mathrm{T}}C^{\mathrm{T}}B_c^{\mathrm{T}})G. \quad (A.28)$$

Using (2.19) to eliminate the terms in the parenthesis, (A.28) is equivalent to

$$0 = AQ + QA^{\mathrm{T}} + V_1 - QC^{\mathrm{T}}B_c^{\mathrm{T}}G - G^{\mathrm{T}}B_c CQ + G^{\mathrm{T}}B_c V_2 B_c^{\mathrm{T}}G. \quad (A.29)$$

Forming $(A.29) + Q\bar{\Sigma}Q - Q\bar{\Sigma}Q$ and noting

$$(QC^{\mathrm{T}}V_2^{-1} - G^{\mathrm{T}}B_c)V_2(QC^{\mathrm{T}}V_2^{-1} - G^{\mathrm{T}}B_c)^{\mathrm{T}} = -QC^{\mathrm{T}}B_c^{\mathrm{T}}G - G^{\mathrm{T}}B_c CQ + G^{\mathrm{T}}B_c V_2 B_c^{\mathrm{T}}G + Q\bar{\Sigma}Q,$$

yields (3.7). Next, computing $(2.17) - (3.7)$, using (A.12b) and noting the identity $Q_{12}C_c^{\mathrm{T}}B^{\mathrm{T}} = G^{\mathrm{T}}Q_2 C_c^{\mathrm{T}}B^{\mathrm{T}} = \hat{Q}P\bar{\Sigma}$, we can obtain (3.9).

Finally, to prove (3.10), using (2.16) and (2.9), (2.14) becomes

$$J(A_c, B_c, C_c) = \mathrm{tr}\, P_1 V_1 + \mathrm{tr}\, P_2 B_c V_2 B_c^{\mathrm{T}}. \quad (A.30)$$

Next, noting the identity $\mathrm{tr}\, P_2 B_c V_2 B_c^{\mathrm{T}} = \mathrm{tr}\, V_2 B_c^{\mathrm{T}}P_2 B_c$, and using (3.1b) and the fact that $(I - P_2 P_2^{\dagger})P_2 = P_2(I - P_2^{\dagger}P_2) = 0$, yields $\mathrm{tr}\, P_2 B_c V_2 B_c^{\mathrm{T}} = \mathrm{tr}\, CQ\Gamma^{\mathrm{T}}P_2 \Gamma QC^{\mathrm{T}}V_2^{-1}$. Thus, using (A.11a), the above expression can be reduced to

$$\mathrm{tr}\, P_2 B_c V_2 B_c^{\mathrm{T}} = \mathrm{tr}\, CQ\hat{P}QC^{\mathrm{T}}V_2^{-1} = \mathrm{tr}\, \hat{P}QC^{\mathrm{T}}V_2^{-1}CQ = \mathrm{tr}\, \hat{P}Q\bar{\Sigma}Q. \quad (A.31)$$

Furthermore, using (A.31) and (A.12a), (A.30) is equivalent to (3.10). Similarly, using the dual approach and noting that $J(A_c, B_c, C_c) = \mathrm{tr}\, \tilde{Q}\tilde{R}$ yields (3.11). $\square$
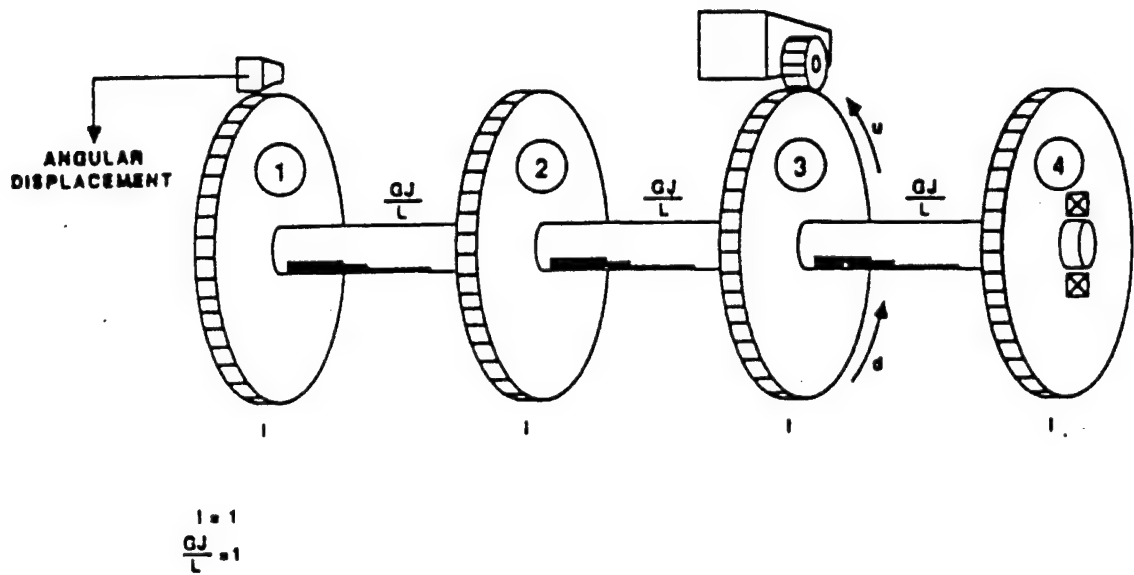
Figure 5.1. The Four Disk Model
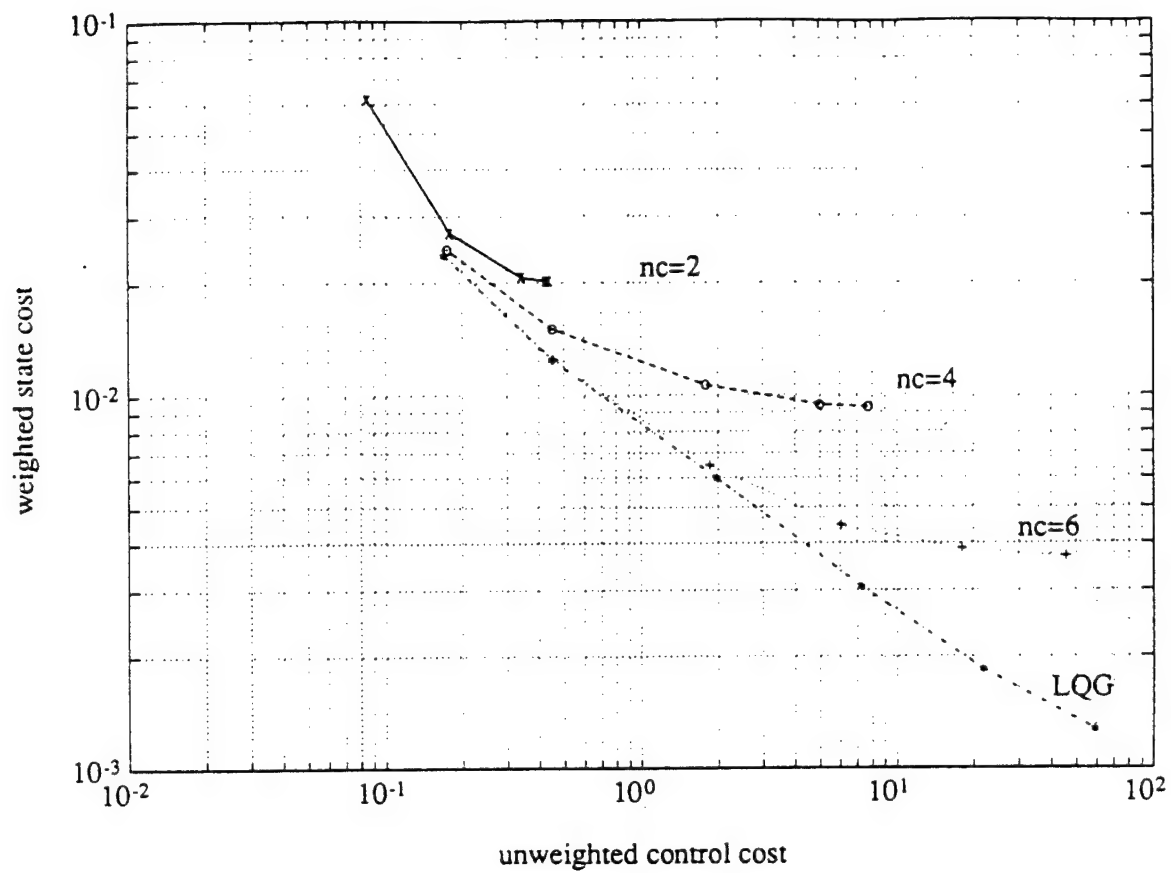
Figure 5.2. Comparison of the Performance Curves for Various Order Controllers for Four Disk Example
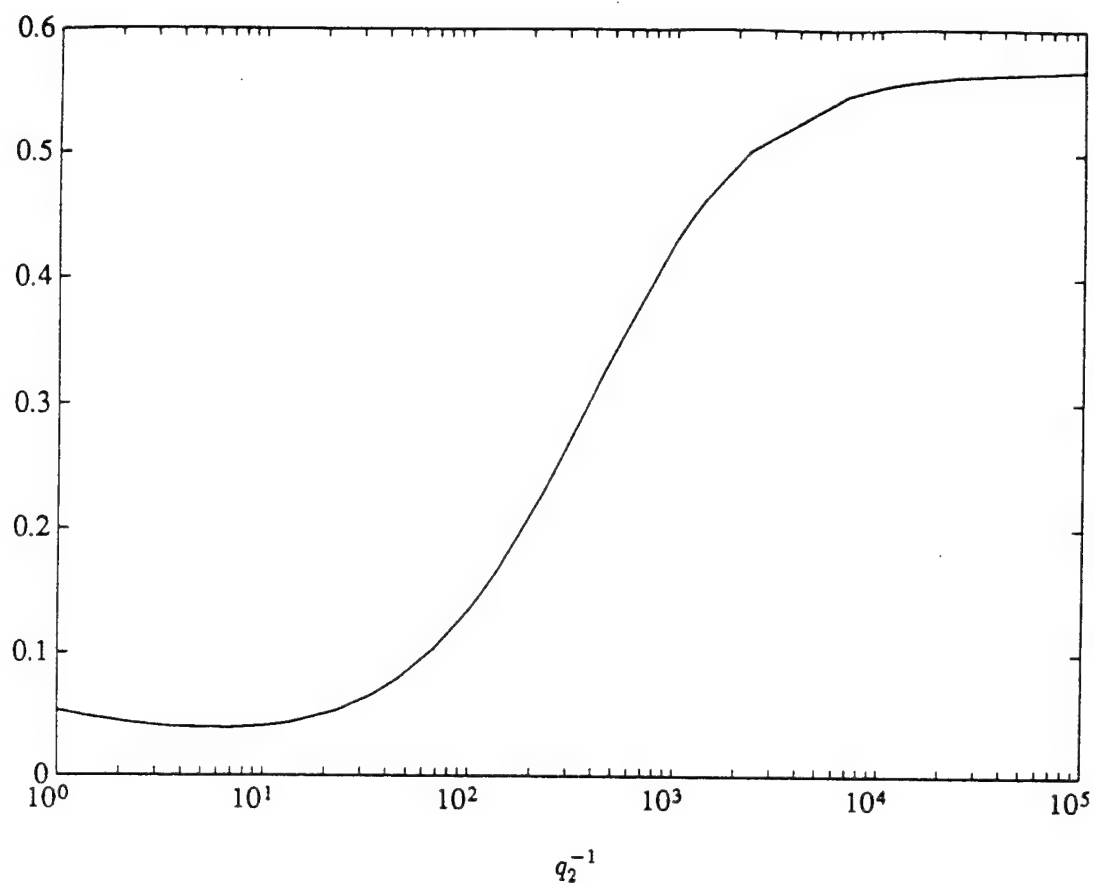
Figure 5.3. $\|\hat{P}\|_F$ as a Function of Control Authority (q_2^{-1}) for Four Disk Example with $n_c = 4$

References

1. M. Athans, *The Role and Use of the Stochastic Linear- Quadratic-Gaussian Problem in Control System Design*, IEEE Trans. Autom. Contr., AC-16(1971), pp. 529–552.

2. B.D.O. Anderson and J.B. Moore, *Linear Optimal Control*, Prentice-Hall, Englewood Cliffs, N.J., 1971.

3. H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*, John Wiley and Sons, New York, 1972.

4. M. Athans, *The Matrix Minimum Principle*, Inform. Control, 11(1968), pp. 592–606.

5. W.S. Levine, T.L. Johnson, M. Athans, *Optimal Limited State Variable Feedback Controllers for Linear Systems*, IEEE Trans. Autom. Contr., AC-16(1971), pp. 785–793.

6. D. C. Hyland, and D. S. Bernstein, *The Optimal Projection Equations for Fixed-Order Dynamic Compensation*, IEEE Trans. Autom. Contr., AC-29(1984), pp. 1034–1037.

7. W. M. Haddad, *Robust Optimal Projection Control-System Synthesis*, Ph.D. Dissertation, Florida Institute of Technology, Melbourne, FL 1987.

8. D. S. Bernstein and W. M. Haddad, *The Optimal Projection Equations with Petersen-Hollot Bounds: Robust Stability and Performance via Fixed-Order Dynamic Compensation for Systems with Structured Real-Valued Parameter Uncertainty*, IEEE Trans. Autom. Contr., 33 (1988), pp. 578–582.

9. D. S. Bernstein and W. M. Haddad, *LQG Control with An H_∞ Performance Bound: A Riccati Equation Approach*, IEEE Trans. Autom. Contr., 34(1989), pp. 293–305.

10. A. Yousuff and R.E. Skelton, *A Note on Balanced Controller Reduction*, IEEE Trans. Autom. Contr., AC-29(1984), pp.254–257.

11. R.E. Skelton, *Dynamic Systems Control*, John Wiley and Sons, New York, 1988.

12. M. Mercadal, *H_2 Fixed Architecture Control Design for Large Scale Systems*. Ph.D. Dissertation, Dept. of Aeronautics and Astronautics, MIT, June 1990.

13. E.G. Collins, Jr., W.M. Haddad, and S.S. Ying, *Reduced-Order Dynamic Compensation using the Hyland-Bernstein Optimal Projection Equations*, AIAA J. Guid. Contr. Dyn., submitted.

14. S.S. Ying, *Reduced-Order H_2 Modeling and Control Using the Optimal Projection Equations: Theoretical Issues and Computational Algorithms*, Ph.D. Dissertation, Florida Institute of Technology, Melbourne, FL, 1993.

15. C.R. Rao and S.K. Mitra, *Generalized Inverse of Matrices and its Applications*, John Wiley and Sons, New York, 1971.

16. A. Albert, *Conditions for Positive and Nonnegative Definiteness in Terms of Pseudo Inverse*, SIAM J. Appl. Math., 17(1969), pp.434-440.

17. D.S. Bernstein and W.M. Haddad, *Robust Stability and Performance via Fixed-Order Dynamic Compensation with Guaranteed Cost Bounds*, Math. Control Signal Systems. 3(1990). pp. 139–163.

18. K. Glover, *All Optimal Hankel Norm Approximations of Linear Multivariable Systems and Their L^∞-error Bounds*, Int. J. Contr., 39(1984), pp. 1115–1193.

19. W.M. Wonham, *Linear Multivariable Control: A Geometric Approach*, Springer-Verlag, New York, 1979.

20. E.G. Collins, Jr., W.M. Haddad, and S.S. Ying, *Construction of Low Authority, Nearly Non-Minimal LQG Compensators for Reduced-Order Control Design*, American Control Conference, 1994, to be published.

21. R. H. Cannon and D.E. Rosenthal, *Experiments in Control of Flexible Structures with Non-colocated Sensors and Actuators*, AIAA Journal of Guidance, Control and Dynamics, 7(1984), pp. 546-553.

22. B.D.O. Anderson and Y. Liu, *Controller Reduction: Concepts and Approaches*, IEEE Transactions on Automatic Control, 34(1989), pp. 802–812.

23. Y. Liu, B.D.O. Anderson, and U-L Ly, *Coprime Factorization Controller Reduction with Bezout Identity Induced Frequency Weighting*, Automatica, 26(1990), pp. 233-249.

24. D.C. Hyland and S. Richter, *On Direct Versus Indirect Methods for Reduced-order Controller Design*, IEEE Transactions on Automatic Control, 35(1990), pp. 377-379.

25. J.W. Brewer, *Kronecker Products and Matrix Calculus in System Theory*, IEEE Trans. Circuit and Systems, 25(1978), pp 772–781.

Appendix I:

Construction of Low Authority, Nearly Non-Minimal LQG Compensators
for Reduced-Order Control Design

Construction of Low Authority, Nearly Non-Minimal
LQG Compensators for Reduced-Order Control Design

Emmanuel G. Collins, Jr.
Harris Corporation
Government Aerospace Systems Div.
MS 22/4849
Melbourne, FL 32902
(407) 727-6358
ecollins@x102a.ess.harris.com

Wassim M. Haddad
School of Aerospace Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0150
(404) 894-2760
wassim.haddad@aerospace.gatech.edu

Sidney S. Ying
Collins Commercial Avionics Division
Rockwell International Corporation
MS 306-100
Melbourne, FL 32934
(407) 768-7063
ssy@dllws.cca.cr.rockwell.com

Abstract

It has been observed numerically that suboptimal
controller reduction methods tend to work well when
applied to low authority LQG controllers. However, to
date, a rigorous justification for this phenomena has not
been established. This paper shows that for continuous-
time stable systems, by proper choice of the structure
of the design weights, the corresponding LQG compen-
sator becomes nonminimal as the control authority is
decreased. An example illustrates that the near non-
minimality of the LQG compensator can result in near
optimality of the corresponding controller obtained by
suboptimal controller reduction.

1. Introduction

The development of linear-quadratic-gaussian (LQG)
theory was a major breakthrough in modern control the-
ory since it provides a systematic way to synthesize high
performance controllers for nominal models of complex,
multi-input multi-output systems. However, one of the
well known deficiencies of an LQG compensator is that
its minimal dimension is usually equal to the dimension
of the design plant. This has led to the development of
techniques to synthesize reduced-order approximations
of the optimal full-order compensator (i.e., controller re-
duction methods) [1-6].

The controller reduction methods almost always
yield suboptimal (and sometimes destabilizing) reduced-
order control laws since an optimal reduced-order con-
troller is not usually a direct function of the parameters
used to compute or describe the optimal full-order con-
troller. Nevertheless, these methods are computation-
ally inexpensive and sometimes do yield high performing
and even nearly optimal control laws. An observation
that holds true about most of these methods is that they
tend to work best at low control authority [4, 6]. How-
ever, to date no rigorous explanation has been presented
to explain this phenomenon.

This paper provides a constructive way of choosing
the weights in a LQG control problem of dimension n
such that for a given $n_c < n$ the corresponding n_c^{th}-order
controller obtained by a suboptimal reduction method is
guaranteed to have essentially the same performance as
the LQG controller at low control authority. Although
the guarantee is for a low authority control problem, it
is expected that, as the control authority is increased
by scaling the appropriate weights, suboptimal reduc-
tion methods will perform better than they would with
another set of weights.

The discussion here focuses on stable systems. It
is shown that if the state weighting matrix R_1 or dis-
turbance intensity V_1 has a specific structure in a basis
in which the A matrix is upper or lower block triangu-
lar, respectively, then at low control authority the corre-
sponding LQG compensator is nearly nonminimal with
minimal dimension n_c. It follows that the LQG com-
pensator can be easily reduced to a n_c^{th}-order controller
having nearly the same performance.

A special case of the conditions presented for R_1
and V_1 has a strong physical interpretation for structural
control problems. In particular, assume that all of the
eigenvalues in the plant are complex and that n_c is an
even number. Then, either R_1 is allowed to weight only
$n_c/2$ modes or V_1 is allowed to disturb only $n_c/2$ modes.

Notation

$\mathbb{R}, \mathbb{R}^{r \times s}, \mathbb{R}^r$	real numbers, $r \times s$ real matrices, $\mathbb{R}^{r \times 1}$
$\mathbb{E}$	expected value
$X \geq 0$	matrix X is nonnegative definite
$X > 0$	matrix X is positive definite
$0_{r \times s}, 0_r$	$r \times s$ zero matrix, $r \times r$ zero matrix
I_r	$r \times r$ identity matrix
$\text{vec}(\cdot)$	the invertible linear operator defined as $\text{vec } S \triangleq [s_1^T \ s_2^T \cdots s_q^T]^T$, $S \in \mathbb{R}^{p \times q}$, where $s_j \in \mathbb{R}^p$ is the jth column of S.

2. Low Authority LQG Compensation

Consider the nth-order linear time-invariant plant

$$\dot{x}(t) = Ax(t) + Bu(t) + D_1 w(t), \qquad (2.1a)$$

$$y(t) = Cx(t) + D_2 w(t), \qquad (2.1b)$$

where (A, B) is stabilizable, (A, C) is detectable, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $y \in \mathbb{R}^l$, and $w \in \mathbb{R}^d$ is a standard white
noise disturbance with intensity I_d and rank $D_2 = l$.
The intensities of $D_1 w(t)$ and $D_2 w(t)$ are thus given,
respectively, by $V_1 \triangleq D_1 D_1^T \geq 0$, and $V_2 \triangleq D_2 D_2^T > 0$.
For convenience, we assume that $V_{12} \triangleq D_1 D_2^T = 0$,
i.e., the plant disturbance and measurement noise are
uncorrelated. Then, the LQG compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \qquad (2.2a)$$

$$u(t) = -C_c x_c(t), \qquad (2.2b)$$

for the plant (2.1) minimizing the steady-state quadratic
performance criterion

$$J(A_c, B_c, C_c) \triangleq \lim_{t \to \infty} \frac{1}{t} \mathbb{E} \int_0^t [x^T(s)R_1 x(s) + u^T(s)R_2 u(s)]ds.$$

$$(2.3)$$

where $R_1 \geq 0$ and $R_2 > 0$ are the weighting matrices for the controlled states and controller input, respectively, is given by:

$$A_c = A - \Sigma P - Q\bar{\Sigma}, \qquad (2.4a)$$

$$B_c = QC^T V_2^{-1}, \quad C_c = R_2^{-1} B^T P, \qquad (2.4b,c)$$

where $\Sigma \triangleq BR_2^{-1}B^T$, $\bar{\Sigma} \triangleq C^T V_2^{-1} C$, and P and Q are the unique, nonnegative-definite solutions respectively of

$$0 = A^T P + PA + R_1 - P\Sigma P. \qquad (2.5)$$

$$0 = AQ + QA^T + V_1 - Q\bar{\Sigma}Q. \qquad (2.6)$$

Furthermore, the "shifted" observability and controllability gramians [1] of the compensator, $\hat{P}$ and $\hat{Q}$, are the unique, nonnegative-definite solutions respectively of

$$0 = (A - Q\bar{\Sigma})^T \hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P, \quad (2.7)$$

$$0 = (A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^T + Q\bar{\Sigma}Q. \quad (2.8)$$

Although a cross-weighting term of the form $2x^T(t)R_{12} u(t)$ can also be included in (2.3), we shall not do so here to facilitate the presentation. The magnitudes of R_2 and V_2 relative to the state weighting matrix R_1 and plant disturbance intensity V_1 govern the regulator and estimator authorities, respectively. The selection of R_2 and V_2 such that $\|R_2\| >> \|R_1\|$, or $\|V_2\| >> \|V_1\|$, yields a low authority compensator. This section shows that when the open-loop plant is stable and (A, R_1) or (A, V_1) have a particular structure, the LQG controller approaches nonminimality as the controller authority decreases. In order to prove this result, we first exploit some interesting structural properties of the solutions of the Riccati equations and Lyapunov equations assuming the coefficient matrix A and the constant driving term R_1 have certain partitioned forms.

Lemma 2.1. Suppose

$$A = \begin{bmatrix} A_1 & 0 \\ A_{21} & A_2 \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad R_1 = \begin{bmatrix} R_{1,1} & 0 \\ 0 & 0_{n-n_r} \end{bmatrix},$$
$$(2.9a,b,c)$$
where $A_1, R_{1,1} \in \mathbb{R}^{n_r \times n_r}$, $B_1 \in \mathbb{R}^{n_r \times m}$, $R_{1,1} > 0$.

(i) If both (A, B) and (A_1, B_1) are stabilizable, then the unique, nonnegative-definite solution of the Riccati equation:

$$0 = A^T P + PA + R_1 - PBB^T P, \qquad (2.10)$$

is given by

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \qquad (2.11)$$

where the $n_r \times n_r$ matrix P_1 is the unique, positive-definite solution of

$$0 = A_1^T P_1 + P_1 A_1 + R_{1,1} - P_1 B_1 B_1^T P_1. \quad (2.12)$$

(ii) If A is asymptotically stable, then the unique, nonnegative-definite solution of the Lyapunov equation:

$$0 = A^T P + PA + R_1, \qquad (2.13)$$

is given by

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}, \qquad (2.14)$$

where the $n_r \times n_r$ matrix P_1 is the unique, positive-definite solution of

$$0 = A_1^T P_1 + P_1 A_1 + R_{1,1}. \qquad (2.15)$$

Proof.

(i) Since (A, B) is stabilizable and $R_1 \geq 0$, it follows from Theorem 12.2 of [7] that there exists a unique, nonnegative-definite solution of the Riccati equation (2.10). Similarly, the assumptions that (A_1, B_1) is stabilizable and $R_{1,1} > 0$ imply that there exists a positive-definite matrix P_1 satisfying the Riccati equation (2.12). Using (2.12), it follows by construction that (2.11) is the solution of (2.10).

(ii) This is a special case of the Riccati equation of property (i). $\square$

The following lemma states the dual of Lemma 2.1 if the coefficient matrix A is upper block triangular and V_1 is upper block diagonal.

Lemma 2.2. Suppose

$$A = \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix}, \quad C = [C_1 \quad C_2], \quad V_1 = \begin{bmatrix} V_{1,1} & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}.$$
$$(2.16a,b,c)$$
where $A_1, V_{1,1} \in \mathbb{R}^{n_r \times n_r}$, $C_1 \in \mathbb{R}^{l \times n_r}$, $V_{1,1} > 0$.

(i) If (A, C) and (A_1, C_1) are detectable, then the unique, nonnegative-definite solution of the following Riccati equation:

$$0 = AQ + QA^T + V_1 - QC^T CQ, \qquad (2.17)$$

is given by

$$Q = \begin{bmatrix} Q_1 & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}. \qquad (2.18)$$

where the $n_r \times n_r$ matrix Q_1 is the unique, positive-definite solution of

$$0 = A_1 Q_1 + Q_1 A_1^T + V_{1,1} - Q_1 C_1^T C_1 Q_1. \quad (2.19)$$

(ii) If A is asymptotically stable, then the unique, nonnegative-definite solution of the Lyapunov equation:

$$0 = AQ + QA^T + V_1, \qquad (2.20)$$

is given by

$$Q = \begin{bmatrix} Q_1 & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}. \qquad (2.21)$$

where the $n_r \times n_r$ matrix Q_1 is the unique, positive-definite solution of

$$0 = A_1 Q_1 + Q_1 A_1^T + V_{1,1}. \qquad (2.22)$$

Proof. The proof is dual to the proof of Lemma 2.1. $\square$

The following theorem proves that with proper choice of the weighting matrices, a low authority LQG controller for a stable plant is nearly nonminimal. The proof of this theorem relies on the above two lemmas.

Theorem 2.1. Consider the plant given by (2.1).

(i) Suppose

$$A = \begin{bmatrix} A_1 & 0 \\ A_{21} & A_2 \end{bmatrix}, \quad R_1 = \begin{bmatrix} R_{1,1} & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}. \qquad (2.23a, b)$$

where $A_1, R_{1,1} \in \mathbb{R}^{n_r \times n_r}$, $R_{1,1} > 0$, and A is asymptotically stable. Let

$$V_2 \triangleq \beta V_2 \qquad (2.24)$$

where V_2 is some finite, positive-definite matrix and $\beta \in \mathbb{R}$ is a positive scalar. Then for any $\delta > 0$, there exists N such that for all $\beta > N$,

$$(\lambda_{n_r+1}/\lambda_{n_r}) < \delta. \qquad (2.25)$$

where λ_i represents the ith eigenvalue of $Q\dot{P}$, $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_i \geq \lambda_{i+1}... \geq 0$, and Q and P are the shifted controllability and observability gramians of the corresponding LQG compensator, satisfying (2.8) and (2.7), respectively.

(ii) Suppose

$$A = \begin{bmatrix} A_1 & A_{12} \\ 0 & A_2 \end{bmatrix}, \quad V_1 = \begin{bmatrix} V_{1,1} & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}. \qquad (2.26a, b)$$

where $A_1, V_{1,1} \in \mathbb{R}^{n_r \times n_r}$, $V_{1,1} > 0$, and A is asymptotically stable. Let

$$R_2 \triangleq \alpha \dot{R}_2, \qquad (2.27)$$

where R_2 is some finite, positive-definite matrix and $\alpha \in \mathbb{R}$ is a positive scalar. Then for any $\delta > 0$, there exists N such that for all $\alpha > N$,

$$(\lambda_{n_r+1}/\lambda_{n_r}) < \delta, \qquad (2.28)$$

where λ_i represents the ith eigenvalue of $Q\dot{P}$ and $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_i \geq \lambda_{i+1}... \geq 0$, and

Q and P are the shifted controllability and observability gramians of the corresponding LQG compensator, satisfying (2.8) and (2.7), respectively.

Proof.

(i) Partition $B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \Sigma_1 & \Sigma_{12} \\ \Sigma_{12}^T & \Sigma_2 \end{bmatrix}$ conformal to A in (2.23). The assumptions (2.23) and that A is asymptotically stable imply that (A, B) and (A_1, B_1) are both stabilizable. Thus, it follows from property (i) of Lemma 2.1 that the unique, nonnegative-definite solution P of the Riccati equation (2.5) has the structure given by (2.11), which implies that

$$P\Sigma P = \begin{bmatrix} P_1 \Sigma_1 P_1 & 0 \\ 0 & 0 \end{bmatrix}. \qquad (2.29)$$

Thus, noting the special partitioned structures in (2.29) and (2.23), and that A is asymptotically stable, it follows from property (ii) of Lemma 2.1 that there exists

$$\dot{P}_0 \triangleq \begin{bmatrix} \dot{P}_1 & 0 \\ 0 & 0_{n-n_r} \end{bmatrix}. \qquad (2.30)$$

which is the unique, nonnegative-definite solution of

$$0 = A^T \dot{P}_0 + \dot{P}_0 A + P\Sigma P. \qquad (2.31)$$

where $n_r \times n_r$ matrix $\dot{P}_1$ is the unique, nonnegative definite solution of $0 = A_1^T P_1 + P_1 A_1 + P_1 \Sigma_1 P_1$. Next, computing $(2.31) - (2.7)$ and using (2.24), yields the following modified Lyapunov equation:

$$0 = A^T \Delta\dot{P} + \Delta\dot{P}A + \beta^{-1}[(\bar{\Sigma}_0 QP) + (\bar{\Sigma}_0 QP)^T] \qquad (2.32)$$

where $\bar{\Sigma}_0 \triangleq C^T V_2^{-1} C. \qquad (2.33a)$

$$\Delta\dot{P} \triangleq \dot{P}_0 - P. \qquad (2.33b)$$

Since A is asymptotically stable and Q and P satisfy (2.6) and (2.7), respectively, Q and P are bounded for all β. Next, we rewrite (2.33b) as

$$P = P_0 - \beta^{-1} \Delta P_0. \qquad (2.34)$$

where $\Delta\dot{P}_0$ is the solution of $0 = A^T \Delta\dot{P}_0 + \Delta\dot{P}_0 A + (\bar{\Sigma}_0 QP + (\bar{\Sigma}_0 QP)^T)$. Now, rewriting (2.8) as $0 = (A - \Sigma P)Q + Q(A - \Sigma P)^T + \beta^{-1} Q\bar{\Sigma}_0 Q$, it follows that

$$Q = \beta^{-1} Q_0. \qquad (2.35)$$

where Q_0 satisfies $0 = (A - \Sigma P)Q_0 + Q_0(A - \Sigma P)^T + Q\Sigma_0 Q$. Next, using (2.34) and (2.35), we obtain (for large β)

$$S \triangleq QP = \begin{bmatrix} \beta^{-1}S_1 & \beta^{-1}S_{12} \\ \beta^{-2}S_{21} & \beta^{-2}S_2 \end{bmatrix}.$$

where $S_1 \in \mathbb{R}^{n_r \times n_r}$, $S_2 \in \mathbb{R}^{n-n_r \times n-n_r}$, and S_1 is nonsingular. Note that since Q and P are nonnegative-definite, S is semisimple and the eigenvalues of S are real and nonnegative. Hence, the eigenvalue ratio of S is the same as the corresponding eigenvalue ratio of βS. Next, define

$$S' \triangleq \begin{bmatrix} S_1 & S_{12} \\ 0 & 0 \end{bmatrix},$$

and recognize that $\lim_{\beta \to \infty} \beta S = S'$. Noting that the eigenvalues of S' are the collection of n_r eigenvalues of S_1 plus $(n - n_r)$ zero eigenvalues, and since the eigenvalues of a matrix are continuous with respect to the parameters of the matrix, it follows that for any $\epsilon > 0$, there exists N such that for all $\beta > N$, $\lambda_{S_1,i} - \epsilon < \lambda_{\beta S,i} < \lambda_{S_1,i} + \epsilon$, for $i = 1, ..., n_r$ and $\lambda_{\beta S,i} < \epsilon$, for $i = n_r + 1, ..., n$, and $\lambda_{\beta S,i}$ and $\lambda_{S_1,i}$ represent the ith eigenvalue of βS and S_1, respectively, in descending order. Hence, it follows that for any $\delta > 0$, there exists N such that for all $\beta > N$, $\frac{\lambda_{S,n_r+1}}{\lambda_{S,n_r}} < \delta$.

(ii) The proof is dual to the proof of (i). $\square$

Remark 2.1. Theorem 2.1 provides two ways of weighting matrices selection resulting in a nearly nonminimal, low authority LQG compensator for a stable plant. The first approach starts by transforming the plant A into coordinates such that A has the representation as in equation (2.23a) after transformation. Then select the weighting matrix R_1 with the partitioned form as in (2.23b) and with rank $R_1 = n_r$. By decreasing the authority of the compensator, or, equivalently, increasing $\|V_2\|$ or β, the eigenvalue ratio, $\frac{\lambda_{n_r+1}}{\lambda_{n_r}}$ of the LQG compensator decreases and the LQG compensator approaches nonminimality with minimal dimension of n_r. Using a dual approach, with A and V_1 partitioned as in (2.26), by increasing $\|R_2\|$ or α, the resulting LQG compensator approaches nonminimality. However, in the limiting case, as $\alpha \to \infty$ or $\beta \to \infty$ then it follows from (2.7) and (2.8) that $P \to 0$ and $Q \to 0$, respectively.

Remark 2.2. Note that if A is in a modal form, then it satisfies both (2.23a) and (2.26a) of Theorem 2.1. In this case, R_1 given by (2.23b), describes a state weighting matrix in which only the states pertaining to selected modes are weighted. Similarly, V_1 given by (2.26b) describes a disturbance that excites only certain modes. It is not uncommon for these conditions to be satisfied or nearly satisfied in practice.

Remark 2.3. The continuous-time results stated in Lemma 2.1, 2.2 and Theorem 2.1 are readily extended to their discrete-time counterparts as shown in [8].

3. Numerical Illustrative Examples

To illustrate the proper choices of the weighting matrices resulting in a nearly nonminimal, low authority LQG compensator for a stable continuous-time plant, consider a simply supported beam with two collocated sensor/actuator pairs. Assuming the beam has length 2 and that the sensor/actuator pairs are placed at coordinates $a = \frac{55}{172}$, and $b = \frac{46}{13}$, a continuous-time model (2.1) retaining the first five modes is obtained with

$$A = \text{block-diag}\left(\begin{bmatrix} 0 & 1 \\ -1 & -.01 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -16 & -.04 \end{bmatrix}, \right.$$
$$\left. \begin{bmatrix} 0 & 1 \\ -81 & -.09 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -256 & -.16 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ -625 & -.25 \end{bmatrix} \right),$$

$$B = \begin{bmatrix} 0 & 0 \\ -0.2174 & -0.8439 \\ 0 & 0 \\ 0.4245 & -0.9054 \\ 0 & 0 \\ -0.6112 & -0.1275 \\ 0 & 0 \\ 0.7686 & 0.7686 \\ 0 & 0 \\ -0.8893 & 0.9522 \end{bmatrix}, \quad C = B^T.$$

The noise intensities are $V_1 \triangleq D_1 D_1^T = 0.1 I_{10}$ and $V_2 \triangleq D_2 D_2^T = \beta I_2$, and it is assumed that $V_{12} \triangleq D_1 D_2^T = 0$. The design objective is to minimize the continuous-time cost $J = \lim_{t \to \infty} \mathbb{E}[x^T R_1 x + u^T R_2 u]$, where $R_2 = \alpha I_2$. Note that the magnitude of the positive real numbers α and β are the indicators of the controller authority level. For this particular plant, A has the representation as in (2.23a) and (2.26a) with $A_{12} = 0$ and $A_{21} = 0$, respectively. Here, we illustrate the results of property (i) of Theorem 2.1 for the cases of $n_r = 2$ and $n_r = 6$. Setting $\alpha = 0.1$, by selecting the weighting matrix $R_1 = \begin{bmatrix} I_{n_r} & 0 \\ 0 & 0 \end{bmatrix}$, and increasing β (hence, decreasing the compensator authority), the resulting LQG compensator approaches nonminimality with minimal dimension of n_r or, equivalently, $\frac{\lambda_{n_r+1}(\hat{Q}\hat{P})}{\lambda_{n_r}(\hat{Q}\hat{P})} \to 0$ where λ_i is the sorted (in descending order) i^{th} eigenvalue of $\hat{Q}\hat{P}$. Figure 1 shows the ratio curve for $n_r = 2$ with $\beta \in (0.01, 0.1, 1, 10, 10^2, 10^3, 10^4, 10^5, 10^6)$. The curve clearly indicates that the ratio decreases as β increases. To illustrate that suboptimal controller reduction methods yield nearly optimal reduced-order compensators for low authority control problems, Figure 1 also shows the norm of the cost gradient of the 2nd-order controller obtained by balancing. The cost gradient is defined as $[(\text{vec}\frac{\partial J}{\partial A_c})^T \quad (\text{vec}\frac{\partial J}{\partial B_c})^T \quad (\text{vec}\frac{\partial J}{\partial C_c})^T]^T$. The cost gradient curve indicates the balanced controller approaches the optimal reduced-order compensator as β increases, or as the control authority decreases. Figure 2 shows the eigenvalue ratio of the LQG controller for $n_r = 6$ and the norm of the cost gradient of the corresponding 6th-order balanced controller.

Conversely. if the weighting term R_1 for the above example does not have the structure given by (2.23b). decreasing the controller authority (i.e., increasing β) may not yield a nearly nonminimal LQG compensator. As an apparent consequence, the norm of the cost gradient of the corresponding 2nd-order balanced controller does not approach zero as the control authority decreases. This is illustrated in Figure 3 for $n_r = 2$ and $R_1 = I_{10}$. Note that for this particular example, at $\beta = 0.01$ the balanced controller destabilizes the closed-loop system and hence the norm of the cost gradient becomes infinite.

4. Conclusion

By exploiting structural properties of the solutions of the Riccati equations and Lyapunov equations, this paper shows that for continuous-time stable systems, if the coefficient matrix A and driving weighting term R_1 (or V_1) have specific structures, the corresponding LQG compensator becomes nonminimal as the control authority is decreased. As illustrated by the example, this near nonminimality can result in near optimality of a controller obtained by suboptimal controller reduction. Conversely, the example shows that if the structure of the weighting matrices do not satisfy the conditions specified in Theorem 2.1. the resulting LQG compensator is not necessarily nearly minimal even at low control authority. In this case, reduced-order controllers obtained by suboptimal projection methods may not be nearly optimal even at low authority.

References

1. A. Yousuff and R.E. Skelton, "A Note on Balanced Controller Reduction," *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp.254–257, March 1984.

2. A. Yousuff and R.E. Skelton, "Controller Reduction by Component Cost Analysis," *IEEE Trans. Autom. Contr.*, Vol. AC-29, pp. 520–530, June 1984.

3. A. Yousuff and R.E. Skelton, "An Optimal Controller Reduction by Covariance Equivalent Realizations," *IEEE Trans. Autom. Contr.*, Vol. AC-31, pp. 56–60, Jan. 1986.

4. C. De Villemagne and R.E. Skelton, "Controller Reduction using Canonical Interactions," *IEEE Trans. Autom. Contr.*, Vol. AC-33, pp. 740–750, Aug. 1988.

5. B. D. O. Anderson and Y. Liu, "Controller Reduction: Concepts and Approaches," *IEEE Transactions on Automatic Control*, Vol. 34, pp 802–812, 1989.

6. Y. Liu, B. D. O. Anderson, and U-L LY, "Coprime Factorization Controller Reduction with Bezout Identity Induced Frequency Weighting," *Automatica*, Vol. 26, pp. 233–249, 1990.

7. W. M. Wonham, *Linear Multivariable Control: A Geometric Approach*, Springer-Verlag, New York, 1979.

8. S. S. Ying, *Reduced-Order H_2 Modeling and Control Using the Optimal Projection Equations: Theoretical Issues and Computational Algorithms*, PhD. Dissertation. Florida Institute of Technology, Melbourne, FL 1993.
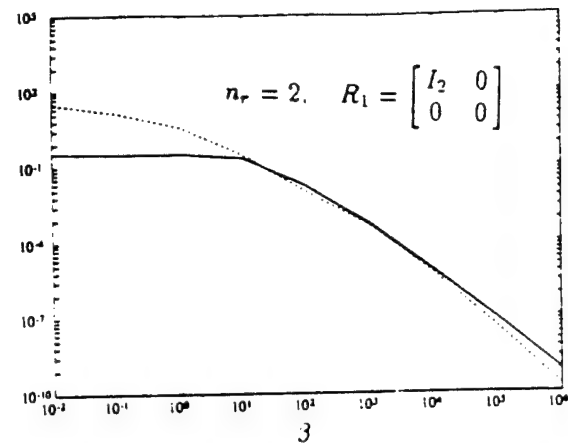
Figure 1. $(\frac{\lambda_{n_r+1}(\hat{Q}\hat{P})}{\lambda_{n_r}(\hat{Q}\hat{P})})$ of the LQG controller (—) and the norm of the cost gradient of the 2nd-order balanced controller (- - -) vs control authority (β) for $n_r = 2$
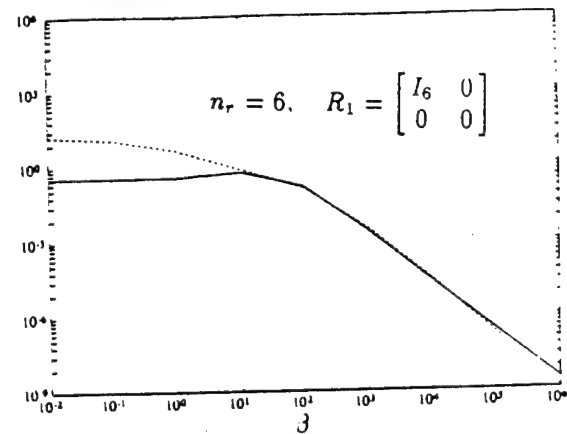
Figure 2. $(\frac{\lambda_{n_r+1}(\hat{Q}\hat{P})}{\lambda_{n_r}(\hat{Q}\hat{P})})$ of the LQG controller (—) and the norm of the cost gradient of the 6th-order balanced controller (- - -) vs control authority (β) for $n_r = 6$
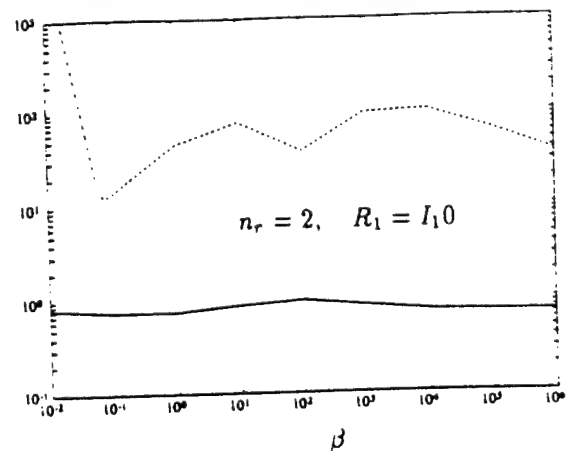
Figure 3. $(\frac{\lambda_{n_r+1}(\hat{Q}\hat{P})}{\lambda_{n_r}(\hat{Q}\hat{P})})$ of the LQG controller (—) and the norm of the cost gradient of the 2nd-order balanced controller (- - -) vs control authority (β) for $n_r = 2$

Appendix J:

An Efficient, Numerically Robust Homotopy Algorithm for H_2 Model Reduction Using the Optimal Projection Equations

An Efficient, Numerically Robust Homotopy Algorithm for H_2 Model Reduction Using the Optimal Projection Equations

by

Emmanuel G. Collins, Jr.
Harris Corporation
Government Aerospace Systems Division
MS. 22/4849
Melbourne, FL 32902
(407) 727-6358
FAX: (407) 727-4016
ecollins@x102a.ess.harris.com

Sidney S. Ying
Rockwell International
Collins Commercial Avionics
MS 306-100
Melbourne, FL 32934
(407) 768-7063
FAX: (407) 254-7805
ssy@dllws.cca.cr.rockwell.com

Wassim M. Haddad
School of Aerospace Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0150
(404) 894-1078
FAX: (404) 894-2760
wassim.haddad@aerospace.gatech.edu

Stephen Richter
Harris Corporation
Government Aerospace Systems Division
MS 22/4849
Melbourne, FL 32902
(407) 727-6358
FAX: (407) 727-6007
srichter@x102a.ess.harris.com

Abstract

Homotopy approaches have previously been developed for synthesizing H_2 optimal reduced-order models. Some of the previous homotopy algorithms were based on directly solving the optimal projection equations, a set of two Lyapunov equations mutually coupled by a nonlinear term involving a projection matrix τ, that characterize the optimal reduced-order model. These algorithms are numerically robust but suffer from the curse of large dimensionality. Subsequently, gradient-based homotopy algorithms were developed. To make these algorithms efficient and to eliminate singularities along the homotopy path, the basis of the reduced-order model was constrained to a minimal parameterization. However, the resultant homotopy algorithms sometimes experienced numerical ill-conditioning or failure due to the minimal parameterization constraint. This paper presents a new homotopy approach to solve the optimal projection equations for H_2 model reduction. The current algorithm avoids the large dimensionality of the previous approaches by efficiently solving a pair of Lyapunov equations coupled by low rank linear operators.

1. Introduction

The continued and pressing need for more accuracy in mathematical modeling of physical processes has led to increasingly high-dimensional models. In order to simplify computer simulations and the design process for feedback compensation, many model reduction schemes have been presented during the last two decades. Among these is the quadratically optimal (or H_2 optimal) model reduction problem. This optimization problem involves determining a reduced-order model of fixed dimension whose outputs approximate the outputs of the original model in a least squares sense. The associated necessary conditions were studied by Wilson in (1970, 1974). Significant simplification of Wilson's results were achieved by recognition and exploitation of an oblique projection matrix by Hyland and Bernstein (1985). The resulting necessary conditions of optimality are characterized by "optimal projection equations" which consist of a pair of $n \times n$ modified Lyapunov equations that are mutually coupled by nonlinear terms involving a projection matrix τ.

The optimal H_2 model order reduction problem is essentially a "younger brother" of the more important problem of optimal H_2 reduced-order controller design. For example, the optimal projection equations for reduced-order modelling are a subset of the optimal projection equations for reduced-order control developed by Hyland and Bernstein (1984). Hence, an important reason for investigating numerical solutions to the model reduction problem is to provide an intermediate step in the development of numerical solutions to the reduced-order control problem.

Several approaches have been considered to synthesize H_2 optimal reduced order models. Based on the first-order necessary condition of optimality, Wilson (1970, 1974) and Hirzinger and Kreisselmeier (1975) presented approaches which implemented the Fletcher-Powell gradient algorithm to minimize the cost over the reduced-order model parameters for a multi-input, multi-output (MIMO) system. Aplevich (1973) and Mishra and Wilson (1980) proposed similar approaches based on the steepest descent algorithm. Using the pole-residue form of the transfer function, Bryson and Carrier (1990) obtained analytical expressions for the first and second order derivatives of the cost function and proposed a Newton-Raphson algorithm for the optimal model reduction of a single-input, single-output (SISO) system. By reformulating the cost function for a SISO system and exploiting its relationship with the coefficients of the transfer function, Spanos *et al.* (1990) developed a two-step gradient-descent algorithm to alternately optimize the numerator and denominator coefficients of the transfer function of the reduced-order model and this algorithm was proved to be globally convergent.

1

Recently, several homotopy algorithms were developed for the H_2 optimal model reduction problem. There are at least three reasons for considering homotopy or continuation methods for optimization problems arising in engineering applications. First of all, it is often desired to find solutions for various values of some set of parameters describing the problem. These parameters can determine the description of the nominal plant, the input authority (in control problems), the amount of system uncertainty, etc... Homotopy methods can be much more efficient in generating these sets of solutions than alternative methods due to the use of prediction steps. (To highlight the importance of the predicition step, various prediction options are illustrated for the homotopy algorithm of this paper via an example.) Secondly, if formulated properly, each intermediate point along a homotopy path has some physical meaning which is useful if the optimization procedure is forced to stop before final convergence. Thirdly, a homotopy path is not a descent path, hence differentiating homotopy methods from most alternative techniques. For nonconvex problems the quickest path to a solution may not be a descent path and hence a homotopy method may actually have faster convergence.

The first homotopy algorithms for H_2 optimal model order reduction were based on directly solving the corresponding optimal projection equations (Zigić et $al.$ 1992,1993a). These algorithms are numerically robust. However, they suffer from the curse of large dimensionality; that is, the corresponding homotopy parameter vector is very large if the original model is large. Hence, these algorithms are intractable for large scale problems.

The above deficiencies led to the development of homotopy algorithms directly based on the gradient expressions (Ge et $al.$ 1993a, 1993b). In these schemes, the parameter vector p represents the reduced-order model. In order to keep the dimension of p relatively small and to avoid high order singularities along the homotopy path, minimal-order parameterizations of the reduced-order model were considered. Because of the reduction in the number of parameters, the resulting algorithms are often more efficient than the original algorithms based on the optimal projection equations. However, since the assumed parameterization may fail to exist or lead to ill-conditioning related to the insistence on using the minimal number of parameters, these resulting algorithms sometimes fail or have very poor convergence properties. One alternative approach proposed by Ge et $al.$ (1993b) is to develop an algorithm that utilizes several minimal parameter homotopies and is capable of switching to an alternative parameterization if ill-conditioning is encountered with the current parameterization. A second approach is to develop an algorithm based on the optimal projection equations that efficiently exploits some of the inherent structure in the matrix design equations and

2

hence reduces the effective size of the homotopy parameter vector in the spirit of the homotopy algorithm described in Collins *et al.* (1993).

This second approach is pursued in this paper. In particular, in order to compute the homotopy curve tangent vectors and the correction steps, the algorithm described here avoids explicit computation and inversion of the Jacobian of the homotopy map as in the homotopy algorithms of Zigić *et al.* (1992,1993a). (It should be acknowledged that Zigić *et al.* (1992,1993a) does not *exactly* invert the Jacobian of the homotopy map, but it does perform an operation that has an equivalent computational burden.) Instead, the algorithm developed here computes the tangent vectors and corrections by solving two Lyapunov equations mutually coupled by linear operators. These equations are efficiently solved using the results of Richter *et al.* (1993) which exploits the low rank properties of the coupling terms. The resultant computational savings over the computational requirements of the algorithms of Zigić *et al.* (1992,1993a) are roughly equivalent to those obtained by computing a solution to a Lyapunov equation via a matrix method (e.g., Brewer (1978) and Lancaster and Tismenetsky (1985)) versus computing a Lyapunov equation solution via solving the associated linear matrix equation $\mathcal{A}\chi = b$ where χ is a vector representing the independent elements of the solution to the Lyapunov equation.

It should be mentioned that the homotopy algorithms of Zigić *et al.* (1992,1993a) are based on arc length and hence allow for singular Jacobians. Hence, they do not assume that the homotopy curve is monotonic with respect to the homotopy parameter. The algorithm here does assume monotonicity. It appears to be possible to extend the algorithm to relax this assumption by using a technique related to that developed by Zigić *et al.* (1993b). However, this is a subject of future research.

The focus of this paper is on computational efficiency. Rigorously proving the existence of the homotopy path that we formulate is beyond the scope of the current paper but is currently being considered in research being performed at the Virginia Polytechnic Institute and State University by Prof. Layne Watson and his students. However, in our computational experience, the homotopy path has always existed.

The paper is organized as follows. Section 2 presents the optimal projection equations for the H_2 model reduction problem. Section 3 gives a brief synopsis of homotopy methods. Next, Section 4 develops a new homotopy algorithm for optimal model reduction design based on the optimal projection equations. Section 5 illustrates the algorithm with three illustrative examples. Finally,

3

Section 6 presents the conclusions.

Nomenclature

$\mathbb{E}$	expected value
$\mathbf{R}^n, \mathbf{R}^{m \times n}$	$n \times 1$ real vectors, $m \times n$ real matrices
$Y \geq X$	$Y - X$ is nonnegative definite
$Y > X$	$Y - X$ is positive definite
x_{ij} or $X_{i,j}$	(i,j) element of matrix X
$X^\dagger$	*Moore-Penrose generalized inverse* of matrix X (Rao and Mitra 1971)
$X^\#$	*Group inverse* of matrix X (Rao and Mitra 1971)
I_r	$r \times r$ identity matrix
$\mathrm{tr}\, X$	trace of square matrix X
$\|X\|_\mathrm{F},\ \|X\|_\mathrm{A}$	Frobenius norm ($\|X\|_\mathrm{F}^2 \triangleq \mathrm{tr}\, X X^\mathrm{T}$), absolute norm ($\|X\|_\mathrm{A} \triangleq \max_{i,j}|X_{i,j}|$)
$\mathrm{vec}(\cdot)$	the invertible linear operator defined such that $\mathrm{vec}(S) \triangleq [s_1^\mathrm{T}\ s_2^\mathrm{T} \cdots s_q^\mathrm{T}]^\mathrm{T},\ S \in \mathbf{R}^{p \times q}$, where $s_j \in \mathbf{R}^p$ denotes the jth column of S.
$e_m^{(i)}$	the m-dimensional column vector whose i^th element equals one and whose additional elements are zeros.
$X(:,k)$	k^th column of the matrix X (MATLAB notation)
$\mathbf{S}^{n \times n}$	the space of symmetric matrices in $\mathbf{R}^{n \times n}$
$X_1^2\mathbf{S}^{n \times n}$	$\mathbf{S}^{n \times n} \times \mathbf{S}^{n \times n}$

2. The Optimal Projection Equations for H_2 Model Reduction

Consider the nth-order, stable, linear time-invariant plant

$$\dot{x}(t) = Ax(t) + Bw(t), \tag{2.1a}$$

4

$$y(t) = Cx(t), \tag{2.1b}$$

where (A, B, C) is controllable and observable, $x(t) \in \mathbf{R}^n$, $w(t) \in \mathbf{R}^m$ is a white noise process with positive-definite intensity V, and $y(t) \in \mathbf{R}^l$. For a given $n_m < n$, the goal of the *optimal fixed-order model-reduction problem* is to determine an n_m^{th}-order model

$$\dot{x}_m(t) = A_m x_m(t) + B_m w(t), \tag{2.2a}$$

$$y_m(t) = C_m x_m(t), \tag{2.2b}$$

where $x_m(t) \in \mathbf{R}^{n_m}$, $y_m(t) \in \mathbf{R}^l$, which minimizes the steady-state quadratic model error criterion

$$J(A_m, B_m, C_m) \triangleq \lim_{t \to \infty} \mathbb{E}[(y - y_m)^{\mathrm{T}} E^{\mathrm{T}} E (y - y_m)], \tag{2.3}$$

where E is an error weighting matrix and $R \triangleq E^{\mathrm{T}} E$ is positive definite. To guarantee that J is finite, it is assumed that A is asymptotically stable and since the value J is independent of the internal realizations of the reduced-order models, we restrict our attention to the set of reduced-order models, $\mathcal{S}_m^+ \triangleq \{(A_m, B_m, C_m) : A_m \text{ is asymptotically stable}, (A_m, B_m) \text{ is controllable and} (A_m, C_m) \text{ is observable}\}$.

Next, forming the augmented system consisting of (2.1) and (2.2),

$$\dot{\tilde{x}}(t) = \tilde{A} \tilde{x}(t) + \tilde{B} w(t), \tag{2.4a}$$

$$\tilde{y}(t) = \tilde{C} \tilde{x}(t), \tag{2.4b}$$

where

$$\tilde{x}(t) \triangleq \begin{bmatrix} x(t) \\ x_m(t) \end{bmatrix}, \quad \tilde{y}(t) \triangleq y(t) - y_m(t),$$

$$\tilde{A} \triangleq \begin{bmatrix} A & 0 \\ 0 & A_m \end{bmatrix}, \quad \tilde{B} \triangleq \begin{bmatrix} B \\ B_m \end{bmatrix}, \quad \tilde{C} \triangleq [C \quad -C_m],$$

allows the cost (2.3) to be expressed as

$$J(A_m, B_m, C_m) \triangleq \lim_{t \to \infty} \mathbb{E}[\tilde{x}^{\mathrm{T}}(t) \tilde{R} \tilde{x}(t)], \tag{2.5a}$$

where $\tilde{R} \triangleq \tilde{C}^{\mathrm{T}} R \tilde{C}$. Next, introduce the weighted performance variables

$$z(t) \triangleq E(y(t) - y_m(t)) = \tilde{E} \tilde{x}(t),$$

where $\tilde{E} \triangleq [EC \quad -EC_m]$. Also define the transfer function from w to z by

$$\tilde{H}(s) \triangleq \tilde{E}(sI_{\tilde{n}} - \tilde{A})^{-1} \tilde{B},$$

5

where $\tilde{n} \triangleq n + n_m$.

Then, it follows that if the augmented system $\tilde{A}$ is asymptotically stable, (2.5a) is given by

$$J(A_m, B_m, C_m) = \|\tilde{H}(s)\|_2^2 \triangleq \frac{1}{2\pi} \int_{-\infty}^{\infty} \|\tilde{H}(j\omega)\|_F^2 d\omega = \text{tr } \tilde{Q}\tilde{R} = \text{tr } \tilde{P}\tilde{V}, \qquad (2.5b)$$

where $\tilde{Q} \triangleq \lim_{t \to \infty} \mathbb{E}[\tilde{x}(t)\tilde{x}^T(t)]$ is the steady-state augmented system covariance, $\tilde{P}$ is its dual, and $\tilde{V} \triangleq \tilde{B}V\tilde{B}^T$. Furthermore, $\tilde{Q}$ and $\tilde{P}$ satisfy the respective Lyapunov equations

$$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V}, \qquad (2.6a)$$

$$0 = \tilde{A}^T\tilde{P} + \tilde{P}\tilde{A} + \tilde{R}. \qquad (2.6b)$$

Before presenting the main theorem we present a key lemma concerning nonnegative definite matrices and several definitions.

Lemma 2.1. (Bernstein and Haddad, 1990). Suppose $\hat{Q} \in \mathbf{R}^{n \times n}$ and $\hat{P} \in \mathbf{R}^{n \times n}$ are symmetric and nonnegative-definite and rank $\hat{Q}\hat{P} = n_m$. Then, the following statements hold:

(i) The $n \times n$ matrix

$$\tau \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#}, \qquad (2.7)$$

is idempotent, i.e., τ is an oblique projection and rank $\tau = n_m$.

(ii) There exist $G, \Gamma \in \mathbf{R}^{n_m \times n}$ and nonsingular $M \in \mathbf{R}^{n_m \times n_m}$ such that

$$\hat{Q}\hat{P} = G^T M \Gamma, \qquad (2.8)$$

$$\Gamma G^T = I_{n_m}. \qquad (2.9)$$

(iii) Finally, if rank $\hat{Q} = \text{rank } \hat{P} = \text{rank } \hat{Q}\hat{P} = n_m$, there exists a nonsingular transformation $W \in \mathbf{R}^{n \times n}$ such that

$$\hat{Q} = W \begin{bmatrix} \Omega & 0 \\ 0 & 0 \end{bmatrix} W^T, \qquad (2.10a)$$

$$\hat{P} = W^{-T} \begin{bmatrix} \Omega & 0 \\ 0 & 0 \end{bmatrix} W^{-1}, \qquad (2.10b)$$

where $\Omega \in \mathbf{R}^{n_m \times n_m}$ is diagonal and nonsingular.

6

Definition 2.1. A triple (G, M, Γ) satisfying property (ii) of Lemma 2.1 is a *projective factorization* of $\hat{Q}\hat{P}$.

Definition 2.2. A model (A_m, B_m, C_m) is an *extremal* of the optimal fixed-order model-reduction problem if it satisfies the first order necessary conditions of optimality, i.e.,

$$\frac{\partial J}{\partial A_m} = 0, \quad \frac{\partial J}{\partial B_m} = 0, \quad \frac{\partial J}{\partial C_m} = 0,$$

where $J(A_m, B_m, C_m)$ is defined by (2.3).

Definition 2.3. A model (A_m, B_m, C_m) is an *admissible extremal* of the optimal fixed-order model-reduction problem if it is an extremal and is also in $\mathcal{S}_m^+$, i.e., the reduced-order model is asymptotically stable, controllable and observable.

Finally, define the positive-definite controllability and observability Gramians

$$W_c \triangleq \int_0^\infty e^{At} BVB^{\mathrm{T}} e^{A^{\mathrm{T}}t} dt, \tag{2.11}$$

$$W_o \triangleq \int_0^\infty e^{A^{\mathrm{T}}t} C^{\mathrm{T}} RC e^{At} dt, \tag{2.12}$$

which satisfy the dual Lyapunov equations

$$0 = AW_c + W_c A^{\mathrm{T}} + BVB^{\mathrm{T}}, \tag{2.13}$$

$$0 = A^{\mathrm{T}} W_o + W_o A + C^{\mathrm{T}} RC. \tag{2.14}$$

Theorem 2.1. (Hyland and Bernstein, 1985). Suppose (A_m, B_m, C_m) is an admissible extremal of the H_2 optimal model-reduction problem. Then there exist $n \times n$ nonnegative-definite matrices $\hat{Q}, \hat{P}$ such that, for some projective factorization (G, M, Γ) of $\hat{Q}\hat{P}$, A_m, B_m and C_m are given by

$$A_m = \Gamma A G^{\mathrm{T}}, \quad B_m = \Gamma B, \quad C_m = C G^{\mathrm{T}}, \tag{2.15a, b, c}$$

and such that $\hat{Q}, \hat{P}$ satisfy

$$0 = A\hat{Q} + \hat{Q}A^{\mathrm{T}} + BVB^{\mathrm{T}} - \tau_\perp BVB^{\mathrm{T}} \tau_\perp^{\mathrm{T}}, \tag{2.16}$$

$$0 = A^{\mathrm{T}}\hat{P} + \hat{P}A + C^{\mathrm{T}} RC - \tau_\perp^{\mathrm{T}} C^{\mathrm{T}} RC \tau_\perp, \tag{2.17}$$

$$\mathrm{rank}\ \hat{Q} = \mathrm{rank}\ \hat{P} = \mathrm{rank}\ \hat{Q}\hat{P} = n_m, \tag{2.18}$$

where τ is given by (2.7) and $\tau_\perp \triangleq I_n - \tau$. Furthermore, the minimal cost is given by

$$J(A_m, B_m, C_m) = \text{tr}\,[C^T R C(W_c - \hat{Q}] = \text{tr}\,[BVB^T(W_o - \hat{P})]. \tag{2.19}$$

Conversely, if there exist $n \times n$ nonnegative-definite matrices $\hat{Q}$ and $\hat{P}$ satisfying (2.16)–(2.18) then the reduced-order model (A_m, B_m, C_m) given by (2.15) is an extremal of the optimal fixed-order model reduction problem. Furthermore, A_m is asymptotically stable if and only if $(\tilde{A}, \hat{E})$ is detectable. In this case, (A_m, B_m) is controllable and (A_m, C_m) is observable.

Remark 2.1. Partitioning $\tilde{Q}$ and $\hat{P}$ given by (2.6a) and (2.6b), respectively, as

$$\tilde{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^T & Q_2 \end{bmatrix}, Q_1 \in \mathbf{R}^{n \times n}, Q_2 \in \mathbf{R}^{n_m \times n_m}, \tag{2.20}$$

$$\hat{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^T & P_2 \end{bmatrix}, P_1 \in \mathbf{R}^{n \times n}, P_2 \in \mathbf{R}^{n_m \times n_m}, \tag{2.21}$$

it follows from Hyland and Bernstein (1985) that $\hat{Q}$ and $\hat{P}$ given by (2.16) and (2.17) can be expressed as

$$\hat{Q} = Q_{12} Q_2^{-1} Q_{12}^T, \tag{2.22}$$

and

$$\hat{P} = P_{12} P_2^{-1} P_{12}^T, \tag{2.23}$$

respectively.

Theorem 2.1 shows that one can compute an optimal reduced-order model by solving a set of coupled, modified Lyapunov equations (2.16) and (2.17) subject to the rank condition constraints (2.18). One approach to find a solution of (2.16) and (2.17) is based on homotopy methods.

3. Homotopy Methods for the Solution of Nonlinear Algebraic Equations

A "homotopy" is a continuous deformation of one function into another. Over the past several years, homotopy or continuation methods (whose mathematical basis is algebraic topology and differential topology (Lloyd 1978)) have received significant attention in the mathematics literature and have been applied successfully to several important problems (Avila 1974, Wacker 1978, Alexander and Yorke 1978, Garcia and Zangwill 1981, Eaves, *et al.* 1983, Watson, 1986). Recently, the engineering literature has also begun to recognize the utility of these methods for engineering applications (see e.g., Richter and DeCarlo 1983, 1984, Turner and Chun 1984, Dunyak *et al.* 1984, Lefebvre *et al.* 1985, Sebok *et al.* 1986, Horta *et al.* 1986, Kabamba *et al.* 1987, Shin *et al.*

1988, Rakowska *et al.* 1991). The purpose of this section is to provide a very brief description of homotopy methods for finding the solutions of nonlinear algebraic equations. The reader is referred to Watson (1986, 1987), and Richter and DeCarlo (1983) for additional details.

The basic problem is as follows. Given sets U and V contained in $\mathbf{R}^n$ and a mapping $F: U \to V$, find solutions $u \in U$ to satisfy

$$F(u) = 0. \tag{3.1}$$

Homotopy methods embed the problem (3.1) in a larger problem. In particular let $H: U \times [0,1] \to \mathbf{R}^n$ be such that:

1) $H(u, 1) = F(u)$.

2) There exists at least one known $u_0 \in \mathbf{R}^n$ which is a solution to $H(\cdot, 0) = 0$, i.e.,

$$H(u_0, 0) = 0. \tag{3.2}$$

3) There exists a continuous curve $(u(\lambda), \lambda)$ in $\mathbf{R}^n \times [0,1]$ such that

$$H(u(\lambda), \lambda) = 0 \text{ for } \lambda \in [0,1], \tag{3.3}$$

   with

$$(u(0), 0) = (u_0, 0). \tag{3.4}$$

4) The curve $(u(\lambda), \lambda)$ is differentiable.

A homotopy algorithm then constructs a procedure to compute the actual curve such that the initial solution $u(0)$ is transformed to a desired solution $u(1)$ satisfying

$$0 = H(u(1), 1) = F(u(1)). \tag{3.5}$$

Now, differentiating $H(u(\lambda), \lambda) = 0$ with respect to λ yields Davidenko's differential equation

$$\frac{\partial H}{\partial u} \frac{du}{d\lambda} + \frac{\partial H}{\partial \lambda} = 0, \tag{3.6}$$

which together with $u(0) = u_0$, defines an initial value problem. The desired solution $u(1)$ is then obtained by numerical integration from 0 to 1. Some numerical integration schemes are described in Watson (1986, 1987).

9

4. A Homotopy Algorithm for H_2 Optimal Reduced-Order Modeling

This section first introduces a homotopy map based on the optimal projection equations and then presents the linearly coupled Lyapunov equations that must be solved for the prediction and correction steps. Next, the homotopy algorithm for the optimal model reduction problem is presented. Finally, the initialization of the homotopy algorithm is discussed in detail.

4.1. The Homotopy Map

To define the homotopy map we assume that the plant matrices (A, B, C), the error weighting matrix R and the disturbance intensity matrix V are functions of the homotopy parameter $\lambda \in [0, 1]$. In particular, it is assumed that

$$\begin{bmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & 0 \end{bmatrix} = \begin{bmatrix} A_0 & B_0 \\ C_0 & 0 \end{bmatrix} + \lambda \left(\begin{bmatrix} A_f & B_f \\ C_f & 0 \end{bmatrix} - \begin{bmatrix} A_0 & B_0 \\ C_0 & 0 \end{bmatrix} \right), \tag{4.1}$$

$$R(\lambda) = R_0 + \lambda(R_f - R_0), \tag{4.2}$$

$$V(\lambda) = V_0 + \lambda(V_f - V_0). \tag{4.3}$$

Note that the above equations imply that $A(0) = A_0$, $B(0) = B_0$, ..., $V(0) = V_0$, and that $A(1) = A_f$, $B(1) = B_f$, ..., $V(1) = V_f$. For notational simplification, we also define

$$\Sigma(\lambda) \triangleq B(\lambda)V(\lambda)B^{\mathrm{T}}(\lambda), \quad \bar{\Sigma}(\lambda) \triangleq C^{\mathrm{T}}(\lambda)R(\lambda)C(\lambda). \tag{4.4a,b}$$

The homotopy formulation $0 = H((\hat{Q}(\lambda), \hat{P}(\lambda)), \lambda)$ is thus given by

$$0 = A(\lambda)\hat{Q}(\lambda) + \hat{Q}(\lambda)A(\lambda)^{\mathrm{T}} + \tau(\lambda)\Sigma(\lambda) + \Sigma(\lambda)\tau^{\mathrm{T}}(\lambda) - \tau(\lambda)\Sigma(\lambda)\tau^{\mathrm{T}}(\lambda), \tag{4.5}$$

$$0 = A(\lambda)^{\mathrm{T}}\hat{Q}(\lambda) + \hat{Q}(\lambda)A(\lambda) + \tau^{\mathrm{T}}(\lambda)\bar{\Sigma}(\lambda) + \bar{\Sigma}(\lambda)\tau(\lambda) - \tau^{\mathrm{T}}(\lambda)\bar{\Sigma}(\lambda)\tau(\lambda), \tag{4.6}$$

where

$$\operatorname{rank} \hat{Q}(\lambda) = \operatorname{rank} \hat{P}(\lambda) = \operatorname{rank} \hat{Q}(\lambda)\hat{P}(\lambda) = n_m, \tag{4.7}$$

$$\tau(\lambda) = \hat{Q}(\lambda)\hat{P}(\lambda)[\hat{Q}(\lambda)\hat{P}(\lambda)]^{\#}, \tag{4.8}$$

and $\lambda \in [0, 1]$.

4.2. The Derivative and Correction Equations

The homotopy algorithm presented later in this section uses a predictor/corrector numerical integration scheme. The prediction step requires derivatives $(\dot{\hat{Q}}(\lambda), \dot{\hat{P}}(\lambda))$, where $\dot{M} \triangleq \frac{dM}{d\lambda}$, while

10

the correction step is based on using a Newton correction, denoted as $(\Delta\hat{Q}, \Delta\hat{P})$. Before constructing the derivative and correction equations, we state the following useful properties about the contragredient transformation of $(\hat{Q}, \hat{P})$.

Using Lemma 2.1, equations (4.7) and (4.8) imply that

$$\hat{Q}(\lambda) = W(\lambda)\Lambda(\lambda)W^{\mathrm{T}}(\lambda) = W_1(\lambda)\Omega(\lambda)W_1^{\mathrm{T}}(\lambda), \tag{4.9}$$

$$\hat{P}(\lambda) = U^{\mathrm{T}}(\lambda)\Lambda(\lambda)U(\lambda) = U_1(\lambda)\Omega(\lambda)U_1^{\mathrm{T}}(\lambda), \tag{4.10}$$

and

$$\tau(\lambda) = W(\lambda)\begin{bmatrix} I_{n_m} & 0 \\ 0 & 0 \end{bmatrix} U(\lambda) = W_1(\lambda)U_1^{\mathrm{T}}(\lambda), \tag{4.11}$$

where

$$W(\lambda) \triangleq [W_1(\lambda) \quad W_2(\lambda)], \quad W_1(\lambda) \in \mathbf{R}^{n \times n_m}, \quad W_2(\lambda) \in \mathbf{R}^{n \times (n-n_m)},$$

$$U(\lambda) \triangleq \begin{bmatrix} U_1^{\mathrm{T}}(\lambda) \\ U_2^{\mathrm{T}}(\lambda) \end{bmatrix}, \quad U_1(\lambda) \in \mathbf{R}^{n \times n_m}, \quad U_2(\lambda) \in \mathbf{R}^{n \times (n-n_m)},$$

$$U(\lambda) = W^{-1}(\lambda), \tag{4.12a}$$

or, equivalently,

$$U(\lambda)W(\lambda) = I_n, \tag{4.12b}$$

$$\Lambda(\lambda) \triangleq \begin{bmatrix} \Omega(\lambda) & 0 \\ 0 & 0 \end{bmatrix}, \quad \Omega(\lambda) \in \mathbf{R}^{n_m \times n_m}, \tag{4.13}$$

and $\Omega(\lambda)$ is diagonal and positive definite. For notational simplicity, we omit the argument λ in what follows.

The derivative equations, obtained by differentiating (4.5) and (4.6) with respect to λ, are given by

$$0 = A_u\dot{\hat{Q}} + \dot{\hat{Q}}A_v^{\mathrm{T}} + R_{\hat{Q}}(\dot{\hat{Q}}, \dot{\hat{P}}) + R_{\hat{Q}}^{\mathrm{T}}(\dot{\hat{Q}}, \dot{\hat{P}}) + V_{\hat{Q}} + V_{\hat{Q}}^{\mathrm{T}}, \tag{4.14}$$

and

$$0 = A_w^{\mathrm{T}}\dot{\hat{P}} + \dot{\hat{P}}A_w + R_{\hat{P}}(\dot{\hat{Q}}, \dot{\hat{P}}) + R_{\hat{P}}^{\mathrm{T}}(\dot{\hat{Q}}, \dot{\hat{P}}) + V_{\hat{P}} + V_{\hat{P}}^{\mathrm{T}}. \tag{4.15}$$

The correction equations, derived similarly by using the relationship between Newton's method and a particular error homotopy, are given by

$$0 = A_u\Delta\hat{Q} + \Delta\hat{Q}A_u^{\mathrm{T}} + R_{\Delta\hat{Q}}(\Delta\hat{Q}, \Delta\hat{P}) + R_{\Delta\hat{Q}}^{\mathrm{T}}(\Delta\hat{Q}, \Delta\hat{P}) + E_{\hat{Q}}^*, \tag{4.16}$$

and

$$0 = A_w^{\mathrm{T}}\Delta\hat{P} + \Delta\hat{P}A_w + R_{\Delta\hat{P}}(\Delta\hat{Q}, \Delta\hat{P}) + R_{\Delta\hat{P}}^{\mathrm{T}}(\Delta\hat{Q}, \Delta\hat{P}) + E_{\hat{P}}^*. \tag{4.17}$$

11

The detail derivation of (4.14)–(4.17) and the definitions of all the coefficients are described in Appendix A. Comparing (4.16)-(4.17) with (4.14)-(4.15) reveals that the derivative and correction equations are identical in form. Each set of equations consist of two coupled Lyapunov equations. Since these equations are linear, using Kronecker algebra (Brewer 1978) and exploiting the rank condition (2.18) of $\hat{Q}$ and $\hat{P}$, they can be converted to the vector form $\mathcal{A}\chi = b$ where for (4.14)–(4.15) χ is a vector consisting of the independent elements of $\dot{W}_1$, $\dot{U}_1$ and $\dot{\Omega}$. Hence, $\mathcal{A}$ is a $(2nn_m + n_m^2) \times (2nn_m + n_m^2)$ matrix and must be inverted to compute χ. Thus, inversion is hence very computationally intensive for even relatively small problems (e.g., $n = 10, n_m = 6$).

Fortunately, the coupling terms $R_{\hat{Q}}$ and $R_{\hat{P}}$ which are linear functions of $(\dot{W}_1, \dot{U}_1, \dot{\Omega})$ or, equivalently, $(\dot{\hat{Q}}, \dot{\hat{P}})$ in (4.14) and (4.15), respectively, have relatively low ranks. Hence, the technique of Richter *et al.* (1993), which exploits this low rank property, can be used to *efficiently* solve equations (4.14) and (4.15) (or, equivalently, (4.16) and (4.17)). In particular, this solution procedure which is detailed in Appendix B, requires an inversion of a square matrix of dimension $n_m(m + l)$, which is identical to the dimension of the homotopy Jacobian inverted in the minimal parameterization approach (Ge *et al.* 1993a,1993b). Hence, the computational efforts required by the present approach are comparable to that required by a minimal parameter homotopy.

Also, note that if the homotopy path exists, the solution to the coupled Lyapunov equations will be well-posed. Hence, the matrices A_u and A_w in (4.14)-(4.17) will have the property that any two eigenvalues of a given matrix will not sum to zero.

4.3. Overview of the Homotopy Algorithm

Below, we present an outline of the homotopy algorithm. This algorithm describes a predictor/corrector numerical integration scheme. In order to force the rank conditions (2.18) of $\hat{Q}$ and $\hat{P}$ during intermediate steps, we use the following scheme to update $(\hat{Q}, \hat{P})$ along the homotopy path. First, using (A.15)-(A.17) and (A.29)-(A.31) and the algorithms described in Appendix C, the prediction $(\dot{\hat{Q}}, \dot{\hat{P}})$ and correction $(\Delta\hat{Q}, \Delta\hat{P})$ are converted to $(\dot{W}_1, \dot{U}_1, \dot{\Omega})$ and $(\Delta W_1, \Delta U_1, \Delta\Omega)$, respectively. Note that $\dot{\Omega}$ and $\Delta\Omega$ are forced to be $n_m \times n_m$ diagonal matrices with this formulation. Next, we update (W_1, U_1, Ω) with these predictions/corrections. Finally, new $(\hat{Q}, \hat{P})$ are constructed with updated (W_1, U_1, Ω) using (4.9) and (4.10) and the rank conditions (4.7) are maintained.

There are several options to be chosen initially. These options are enumerated before presenting

12

the actual algorithm. Note that each option corresponds to a particular flag being assigned some integer value.

4.3.1. Prediction Scheme Options

Here we use the notation λ_0, λ_{-1}, and λ_1 representing the values of λ at respectively the current point on the homotopy curve, the previous point, and the next point. Also, $\dot{M} \triangleq dM/d\lambda$ and $\theta(\lambda)$ is a vector representation of $(W_1(\lambda), U_1(\lambda), \Omega(\lambda))$.

$pred = 0$. No prediction. This option assumes that $\theta(\lambda_1) = \theta(\lambda_0)$.

$pred = 1$. Linear prediction. This option assumes that $\theta(\lambda_1)$ is predicted using $\dot{\theta}(\lambda_0)$ and $\theta(\lambda_0)$, i.e., $\theta(\lambda_1) = \theta(\lambda_0) + (\lambda_1 - \lambda_0)\dot{\theta}(\lambda_0)$.

$pred = 2$. Cubic spline prediction.

This prediction of $\theta(\lambda_1)$ requires $\theta(\lambda_0), \dot{\theta}(\lambda_0), \theta(\lambda_{-1})$, and $\dot{\theta}(\lambda_{-1})$. In particular,

$$\text{vec}[\theta(\lambda_1)] = a_0 + a_1\lambda_1 + a_2\lambda_1^2 + a_3\lambda_1^3,$$

where a_0, a_1, a_2, and a_3 are computed by solving

$$[a_0 \ a_1 \ a_2 \ a_3] \begin{bmatrix} 1 & 0 & 1 & 0 \\ \lambda_{-1} & 1 & \lambda_0 & 1 \\ \lambda_{-1}^2 & 2\lambda_{-1} & \lambda_0^2 & 2\lambda_0 \\ \lambda_{-1}^3 & 3\lambda_{-1}^2 & \lambda_0^3 & 3\lambda_0^2 \end{bmatrix} = \begin{bmatrix} \text{vec}[\theta(\lambda_{-1})] \\ \text{vec}[\dot{\theta}(\lambda_{-1})] \\ \text{vec}[\theta(\lambda_0)] \\ \text{vec}[\dot{\theta}(\lambda_0)] \end{bmatrix}^{\text{T}}.$$

Note that if $\theta(\lambda_{-1})$ and $\dot{\theta}(\lambda_{-1})$ are not available (as occurs at the initial iteration of the homotopy algorithm), then $\theta(\lambda_1)$ is predicted using linear prediction, i.e.,

$$\theta(\lambda_1) = \theta(\lambda_0) + (\lambda_1 - \lambda_0)\dot{\theta}(\lambda_0).$$

4.3.2. Basis Options for Solving the Coupled Lyapunov Equations

The main computational burden of the algorithm given below is the solution of the coupled Lyapunov equations (4.14) and (4.15) or, (4.16) and (4.17) at each prediction step or correction iteration. Efficient solutions of these equations, as described in Appendix B, makes the algorithm feasible for large scale systems. The most desired solution procedure is based on diagonalizing the coefficient matrices A_u and A_w of the coupled Lyapunov equations. This is usually possible. However, it is also possible that this diagonalization will be intractable for some points along the

homotopy path. A numerical conditioning test is embedded in the program to determine whether the coefficient matrices are truly diagonalizable. If they are not, then the coupled Lyapunov equations are solved using the Schur decomposition. A second option relies exclusively on the Schur decomposition.

basis = 1. A_u and A_w are diagonalized when solving (4.14)–(4.15) or (4.16)–(4.17).

basis = 2. A_u and A_w are in Schur form when solving (4.14)–(4.15) or (4.16)–(4.17).

4.3.3.   Outline of the Homotopy Algorithm

Step 1. Initialize $loop = 0$, $\lambda = 0$, $\Delta\lambda \in (0, 1]$, $S = S_0$, $(\hat{Q}, \hat{P}) = (\hat{Q}_0, \hat{P}_0)$.

Step 2. Let $loop = loop + 1$. If $loop = 1$, then go to Step 4.

Step 3. Advance the homotopy parameter λ and predict the corresponding $\hat{Q}(\lambda)$ and $\hat{P}(\lambda)$ as follows.

3a. Let $\lambda_0 = \lambda$.

3b. Let $\lambda = \lambda_0 + \Delta\lambda$.

3c. If $pred \geq 1$, then perform the next step to compute $\dot{\hat{Q}}(\lambda_0)$ and $\dot{\hat{P}}(\lambda_0)$ according to (4.14) and (4.15). Else, let $\hat{Q}(\lambda) = \hat{Q}(\lambda_0)$ and $\hat{P}(\lambda) = \hat{P}(\lambda_0)$ and go to step (4), i.e., no prediction is performed.

3d. Transform A_u and A_w into suitable matrix form by using the option defined by *basis*, then solve (4.14) and (4.15) as described in Appendix B.

3e. Compute $(\dot{W}_1(\lambda_0), \dot{U}_1(\lambda_0), \dot{\Omega}(\lambda_0))$ from $(\dot{\hat{Q}}(\lambda_0), \dot{\hat{P}}(\lambda_0))$ by using (A.15)–(A.17) and the procedure described in Appendix C.

3f. Predict $(W_1(\lambda), U_1(\lambda), \Omega(\lambda))$ by using the option defined by *pred*.

3g. Compute $(\hat{Q}(\lambda), \hat{P}(\lambda))$ from $(W_1(\lambda), U_1(\lambda), \Omega(\lambda))$ using (4.9) and (4.10).

Step 4. Correct the current approximations $\hat{Q}(\lambda^*)$ and $\hat{P}(\lambda^*)$ as follows.

4a. Compute the errors $(E_{\hat{Q}}^*, E_{\hat{P}}^*)$ in the correction equations (A.24)–(A.25).

4b. Transform A_u and A_w into suitable matrix form by using the option defined by *basis*, then

14

solve (4.16) and (4.17) as described in the Appendix B for $\Delta\hat{Q}$ and $\Delta\hat{P}$.

4c. Compute $(\Delta W_1, \Delta U_1, \Delta\Omega)$, from $(\Delta\hat{Q}, \Delta\hat{P})$ by using (A.29)-(A.31) and the algorithms described in Appendix C.

4d. Let $W_1(\lambda) \longleftarrow W_1(\lambda) + \Delta W_1,\quad U_1(\lambda) \longleftarrow U_1(\lambda) + \Delta U_1,\quad \Omega(\lambda) \longleftarrow \Omega(\lambda) + \Delta\Omega.$

4e. Compute $(\hat{Q}(\lambda), \hat{P}(\lambda))$ from $(W_1(\lambda), U_1(\lambda), \Omega(\lambda))$ using (4.9) and (4.10).

4f. Recompute the errors $(E_{\hat{Q}}^*, E_{\hat{P}}^*)$ in the correction equations (A.24)-(A.25). If the max $(\frac{\|E_{\hat{Q}}^*\|_A}{\|\Sigma(\lambda)\|_A}, \frac{\|E_{\hat{P}}^*\|_A}{\|\Sigma(\lambda)\|_A}) < \delta^*$, where δ^* is some preassigned correction tolerance, then set $\lambda_0 = \lambda$, and adjust next step size $\Delta\lambda$ according to the number of the correction steps required to converge before going to Step 3b. Else, if the number of corrections exceeds a preset limit, reduce $\Delta\lambda$ and go to Step 3b; otherwise, go to Step 4b.

Step 5. If $\lambda = 1$, then stop. Else, go to Step 2.

Note that the algorithm described above allows the step size $(\Delta\lambda)$ to vary dynamically depending on the speed of convergence which is gauged by the number of the correction steps. If the number is small (e.g., ≤ 3), we increase (e.g., double) the previous step size when computing the next step. If it takes many steps to converge (e.g., > 10), or does not converge, the step size is reduced (e.g., in half). δ^* in Step 4f is a preassigned correction error tolerance which can be assigned with two values in the program. One is the intermediate correction error tolerance which is used when $\lambda < 1$. The other value is the final correction error tolerance which is usually smaller and is used when $\lambda = 1$. The choice of the magnitudes of theses tolerances are problem dependent. In general, the intermediate correction tolerance is desired to be reasonably large to speed the homotopy curve following. However, the algorithm may fail to converge if these tolerances are too large. The final correction tolerance is usually small to ensure the accuracy of the final results.

4.4. Initial System Selection

In this subsection, we discuss the importance of the homotopy initialization and some guidelines for choosing the initial system matrices. It is assumed that the designer has supplied a set of system and weighting matrices, $S_f = (A_f, B_f, C_f, R_f, V_f)$ describing the optimization problem whose solution is desired. In addition, it is assumed that the designer has chosen an initial set of related system matrices $S_0 = (A_0, B_0, C_0, R_0, V_0)$ that has an easily obtained $(\hat{Q}_0, \hat{P}_0)$ which is either a solution or a good approximation to the solution of the optimal projection equations

corresponding to the initial system (i.e., (4.5)–(4.8) with $\lambda = 0$).

While in general homotopy methods ease the restriction that the starting point be close to some optimal of the optimization problem, the initial guess does affect the performance of the homotopy algorithm. An algorithm with an initial estimate close to the optimal solution usually converges fast. Furthermore, different initial systems may lead to different results. However, as illustrated by Example 5.1 below, an initial system with lower cost than an alternative initial system will *not* always lead to an optimal reduced-order model with lower cost. Below we describe an initialization approach utilizing component cost analysis in balanced coordinates (Kabamba 1985, Skelton and Kabamba 1986), to select the initial system matrices S_0. A similar approach is presented in Ge *et al.* (1993a, 1993b).

4.4.1. Initialization Algorithm

(i) Perform a balanced transformation (Moore 1981) on the given system: $x_b \triangleq T_b x$ such that the controllability and observability Gramians in the balanced coordinates are given by $W_{c,b} = W_{o,b} = \text{diag}(\sigma_1, \sigma_2, ..., \sigma_n)$, and $\sigma_1 \geq \sigma_2 ... \geq \sigma_n$.

(ii) Denoting the balanced realization by (A_b, B_b, C_b), compute the component cost for the ith-state $x_{b,i}$ in balanced coordinates: $v_{b,i} = \sigma_i C_b^T(:, i) C_b(:, i)$, where $C_b(:, i)$ is the ith column of C_b.

(iii) Perform a permutation transformation on the balanced state vector $x_a = T_a x_b$ such that in the new coordinates, the component cost is sorted in a descending order, or, equivalently, $v_{a,i} \geq v_{a,j}$ for $i < j$, where $v_{a,i}$ is the component cost associated with $x_{a,i}$.

(iv) Denoting (A_a, B_a, C_a) the triple after the above transformation, partition

$$A_a = \begin{bmatrix} A_{a,11} & A_{a,12} \\ A_{a,21} & A_{a,22} \end{bmatrix}, \quad B_a = \begin{bmatrix} B_{a,1} \\ B_{a,2} \end{bmatrix}, \quad C_a = [C_{a,1} \ C_{a,2}].$$

where $A_{a,11} \in \mathbf{R}^{n_m \times n_m}, B_{a,1} \in \mathbf{R}^{n_m \times m}, C_{a,1} \in \mathbf{R}^{l \times n_m}$. Note that the triple $(A_{a,11}, B_{a,1}, C_{a,1})$ is a reduced-order model by using component cost analysis in balanced coordinates.

(v) Choose the initial system matrices in the sorted component cost coordinates as

$$A_{a,0} = \begin{bmatrix} A_{a,11} & 0 \\ 0 & A_{a,22} \end{bmatrix}, \quad B_{a,0} = \begin{bmatrix} B_{a,1} \\ 0 \end{bmatrix}, \quad C_{a,0} = [C_{a,1} \ 0].$$

where $A_{a,11}, A_{a,22}, B_{a,1}$, and $C_{a,1}$ are given in (iv). The initial system triple (A_0, B_0, C_0) is obtained by transforming $(A_{a,0}, B_{a,0}, C_{a,0})$ back into the original coordinates using T_a and

16

T_b. Next, form an augmented system consisting of (A_0, B_0, C_0) and $(A_{a,11}, B_{a,11}, C_{a,11})$ and compute the initial guess $\hat{Q}_0$ and $\hat{P}_0$ using (2.22) and (2.23), respectively. Since (A_0, B_0, C_0) is a nth- order nonminimal model whose minimal realization can be represented by a minimal n_m^{th}-order triple $(A_{a,11}, B_{a,11}, C_{a,11})$, $(\hat{Q}_0, \hat{P}_0)$ is a solution of the optimal projection equations corresponding to the initial system (i.e., (4.5)–(4.8) with $\lambda = 0$).

Remark 4.1. Another option for choosing the initial system is based on the triple (A_b, B_b, C_b) obtained in (ii) which describes the system in balanced coordinates. In particular, partition

$$A_b = \begin{bmatrix} A_{b,11} & A_{b,12} \\ A_{b,21} & A_{b,22} \end{bmatrix}, \quad B_b = \begin{bmatrix} B_{b,1} \\ B_{b,2} \end{bmatrix}, \quad C_b = [C_{b,1} \ C_{b,2}],$$

where $A_{b,11} \in \mathbf{R}^{n_m \times n_m}, B_{b,1} \in \mathbf{R}^{n_m \times m}, C_{b,1} \in \mathbf{R}^{l \times n_m}$. Note that the triple $(A_{b,11}, B_{b,1}, C_{b,1})$ is a reduced-order model by using the balanced reduction method (Moore 1981). Now, follow the procedure stated in (v) to construct (A_0, B_0, C_0) and $(\hat{Q}_0, \hat{P}_0)$ in the original coordinates.

5. Illustrative Numerical Examples

This section contains results and observations obtained on three examples. It is assumed $V = R = I$ for each example and the cost J is computed using (2.5b). Using these examples, we compare different algorithm options. In particular, we desire to compare the the speed of the algorithm with various prediction options and with the two basis options for solving the coupled Lyapunov equations. The comparison are all based on a MATLAB implementation of the algorithm and the program in each case was run on a 386, 40 MHz PC. Unless otherwise stated, the initial system $\mathcal{S}_0$ and the initial estimate $(\hat{Q}_0, \hat{P}_0)$ for all the solutions are determined using the algorithms described in Section 4.4.1.

Example 5.1. (Villemagne and Skelton 1987). The system given by

$$A = \begin{bmatrix} -1 & 3 & 0 \\ -1 & -1 & 1 \\ 4 & -5 & -4 \end{bmatrix}, \quad B = \begin{bmatrix} -2 \\ 2 \\ 4 \end{bmatrix}, \quad C = [1 \ 0 \ 0],$$

is to be reduced to an optimal 1st-order model. We follow the algorithm described in Section 4.2 and perform a suboptimal model reduction using component cost analysis to construct the initial system and initial guess. The corresponding optimal reduced-order model obtained from the homotopy algorithm is

$$A_m = -10.4365, \ B_m = -1.5972, \ C_m = 1.5972.$$

17

This model yields a cost $J = 1.6882$, which has the same value as the cost obtained in Zigić *et al.* (1993a). However, if a different initial system which is based on the balanced reduction as described in Remark 4.1 is chosen, we obtain

$$A_m = -0.2863, \quad B_m = 0.8152, \quad C_m = 0.8152.$$

This model yields a cost $J = 1.2288$ which is smaller than the above cost and is the same as the cost obtained in Ge *et al.* (1993a,1993b). Thus, using this example we demonstrate that different optimal reduced-order models may be obtained if the initial system is different. This phenomenon was also observed by Ge *et al.* (1993a,1993b).

Example 5.2. (Hickin and Sinha 1980). The following plant

$$A = \begin{bmatrix} -6.2036 & 15.0540 & -9.8726 & -376.5800 & 251.3200 & -162.2400 & 66.8270 \\ 0.5300 & -2.0176 & 1.4363 & 0 & 0 & 0 & 0 \\ 16.8460 & 25.0790 & -43.5550 & 0 & 0 & 0 & 0 \\ 377.4000 & -89.4490 & -162.8300 & 57.9980 & -65.5140 & 68.5790 & 157.5700 \\ 0 & 0 & 0 & 107.2500 & -118.0500 & 0 & 0 \\ 0.3699 & -0.1445 & -0.2630 & -0.6472 & 0.4995 & -0.2113 & 0 \\ 0 & 0 & 0 & 0 & 0 & 376.9900 & 0 \end{bmatrix},$$

$$B = \begin{bmatrix} 89.3530 & 0 \\ 376.9900 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0.2113 \\ 0 & 0 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

is to be reduced to an optimal 2nd-order model.

We use this example to illustrate the effects with various prediction options. Table 5.1 shows some of the run time statistics of the program for acquiring the optimal reduced-order model ($n_m = 2$) for this example when the diagonalizing basis was chosen in solving the coupled Lyapunov equations, and various prediction options were used. The table compares the number of floating point operations, the actual run time, the number of predictions and corrections performed, and the minimum and maximum homotopy step sizes for each prediction option. From the table, note that the homotopy path of the option with no prediction advanced extremely slow and the step size was reduced to a value less than 10^{-14}. This inefficiency of the homotopy algorithm with no prediction is explained using Figure 5.1 and 5.2 which show the behavior of $\|\hat{Q}(\lambda)\|_F$ and $\|\hat{P}(\lambda)\|_F$ with respect to λ, respectively, and hence reveal the magnitudes of the changes of $\hat{Q}(\lambda)$ and $\hat{P}(\lambda)$

along the homotopy path. Both figures indicate steep slopes of the $\|\hat{Q}(\lambda)\|_F$ and $\|\hat{P}(\lambda)\|_F$ curves when $\lambda < 0.2$. The option with no prediction actually sets $\hat{Q}(\lambda+\Delta\lambda) = \hat{Q}(\lambda)$ and $\hat{P}(\lambda+\Delta\lambda) = \hat{P}(\lambda)$, which implies that this option predicts the curves of Figures 5.1 and 5.2 advancing horizontally. However, these figures show that when $\lambda < 0.2$ this prediction is very poor. Therefore, as shown in Table 5.1, even with an extremely small step size (10^{-14}), it is difficult for the no-prediction option to advance at certain points along the homotopy path.

Prediction Option	Megaflops	RealTime (sec.)	Predictions/ Corrections	Minimum Step Size	Maximum Step Size
None	> 208	> 3114	> 166	$< 10^{-14}$	0.01
Linear	33.4	488	59	0.01	0.08
Cubic	22.8	341	40	0.01	0.16

Table 5.1. Run-Time Statistics of Example 2 with Intermediate $\delta^* = 5 \cdot 10^{-4}$.

Also, as would be expected, Table 5.1 indicates that the algorithm using the cubic spline prediction is more efficient (by about 50%) than the algorithm implementing the linear prediction option for this case, and both linear and cubic spline predictions are far more efficient than using no prediction at all. The improvement in efficiency with cubic spline prediction increases when the intermediate error tolerance is reduced, because in addition to the current data point and its gradient used by linear prediction, the cubic spline prediction also utilizes the past data point and its gradient along the homotopy path. This additional information becomes more accurate with a tighter error tolerance. The ability to predict along the curve described by the changing parameters is one of the practical benefits of formulating an optimization problem in terms of a homotopy.

It should be noted that both the linear and cubic spline prediction cases used the same final correction tolerance and yielded the same optimal 2nd-order model:

$$A_m = \begin{bmatrix} -0.1997 & 0.5045 \\ -0.5044 & -13.2750 \end{bmatrix}, \quad B_m = \begin{bmatrix} 14.9616 & -0.0454 \\ 18.4615 & 0.3655 \end{bmatrix}, \quad C_m = \begin{bmatrix} -0.0085 & -0.2234 \\ -14.9617 & 18.4638 \end{bmatrix},$$

with a cost $J = 23249.3$. The normalized output difference between the reduced-order model and the original plant, $\frac{\lim_{t \to \infty} \mathbb{E}[(y-y_m)^T R(y-y_m)]}{\lim_{t \to \infty} \mathbb{E}[y^T Ry]}$, is 15.2% for optimal reduced-order model and 35.4% for the suboptimal reduced-order model obtained by balancing or the component cost analysis in balanced coordinates. (For this problem the above two suboptimal reduction methods yield the same reduced-order model.)

Example 5.3. (Collins *et al.* 1991)   The given system is a state space model of the transfer function between a torque actuator and an approximately collocated torsional rate sensor for the ACES structure (Irwin *et al.* 1988). This SISO system is of order $n = 17$ and with the following

plant matrices

$$A = \text{block-diag}(\begin{bmatrix} -0.0251 & 3.8433 \\ -3.8433 & -0.0251 \end{bmatrix}, \begin{bmatrix} -0.0368 & 4.9057 \\ -4.9057 & -0.0368 \end{bmatrix}, \begin{bmatrix} -0.0485 & 8.9654 \\ -8.9654 & -0.0485 \end{bmatrix},$$

$$\begin{bmatrix} -0.0649 & 12.0770 \\ -12.0770 & -0.0649 \end{bmatrix}, \begin{bmatrix} -0.4281 & 14.6984 \\ -14.6984 & -0.4281 \end{bmatrix}, \begin{bmatrix} -0.1351 & 15.4179 \\ -15.4179 & -0.1351 \end{bmatrix},$$

$$\begin{bmatrix} -5.1522 & 51.4577 \\ -51.4577 & -5.1522 \end{bmatrix}, \begin{bmatrix} -0.0320 & 73.5133 \\ -73.5133 & -0.0320 \end{bmatrix}, -92.3998),$$

$$B^{\mathrm{T}} = [0.0017 \ 0.0436 \ -0.0031 \ 0.0178 \ 0.0117 \ 0.0415 \ -0.0162 \ 0.0148 \ -0.0529$$

$$-0.2765 \ -0.2988 \ -0.0188 \ 0.4444 \ 1.7120 \ -1.6142 \ -2.6348 \ -4.8879 \cdot 10^{-4}],$$

$$C = [-0.0177 \ 0.0266 \ 0.0097 \ 0.0878 \ -0.0057 \ 0.0133 \ -0.0152 \ 0.0264 \ 0.0037$$

$$-0.0090 \ -0.0051 \ 0.0181 \ 0.0165 \ 0.0052 \ -0.0077 \ 0.0026 \ 184.7996].$$

An optimal reduced-order model with $n_m = 6$ is obtained

$$A_m = \begin{bmatrix} -0.0386 & 73.5136 & 0.0083 & 0.0227 & -0.0381 & -0.0303 \\ -73.5136 & -0.0253 & -0.0059 & -0.0212 & 0.0345 & 0.0220 \\ -0.0083 & -0.0059 & -0.0055 & -3.8545 & 0.2623 & 0.0228 \\ 0.0227 & 0.0212 & 3.8545 & -0.0541 & 0.1001 & 0.5725 \\ -0.0381 & -0.0345 & -0.2623 & 0.1001 & -0.1879 & -15.3794 \\ 0.0303 & 0.0220 & 0.0228 & -0.5725 & 15.3794 & -0.0970 \end{bmatrix},$$

$$B_m = \begin{bmatrix} -0.1229 \\ -0.0995 \\ -0.0123 \\ 0.0386 \\ -0.0640 \\ 0.0456 \end{bmatrix}, \quad C_m = \begin{bmatrix} -0.1229 \\ 0.0995 \\ 0.0123 \\ 0.0386 \\ -0.0640 \\ -0.0456 \end{bmatrix}^{\mathrm{T}}.$$

This model yields a cost $J = 6.9521 \cdot 10^{-5}$.

Table 5.2 shows a comparison of the algorithms for solving the optimal reduced-order model ($n_m = 6$) for this example when the linear prediction is chosen and various basis options were used in solving the coupled Lyapunov equations. The results indicate that the diagonalizing option saves about 50% of computation time over the Schur-form option for this particular example.

Basis Option	Megaflops	RealTime (sec.)	Predictions & Corrections
Diagonal	145	897	14
Schur Form	310.5	1343	14

Table 5.2. Run-Time of Example 5.3 using Different Basis in Solving Coupled Lyapunov Equations.

As noted in Zigić *et al.* (1993a), it took 77 hours to solve for an optimal reduced-order model for this example using a DEC work station. The efficient solver for the coupled modified Lyapunov equations and the selection of the initial estimates near the optimal solution significantly reduce the computational time (to about 15 to 25 minutes on a 386, 40 MHz PC, depending on the basis option). In particular, to solve for the 6th-order optimal model which has $n = 17$, $n_m = 6$, $l = m = 1$, the approach described in Appendix B requires solving $m_* = n_m(m + l) = 12$ Lyapunov equations and inverting a 12×12 matrix for each prediction or correction step which averaged about 13.7 Megaflops operations (Schur form basis). On the other hand, the approach proposed in Zigić *et al.* (1993a) involves the computation of the kernel of a Jacobian matrix for each tangent vector computation. The Jacobian is constructed by exploiting the rank condition (2.18) and has a dimension of $2nn_m + n_m^2$ rows and $2n_m + n_m^2 + 1$ or, a 240×241 matrix for this example. The kernel is found by computing a QR factorization of the Jacobian matrix then using a back substitution. By simulation, a typical flop counts for the MATLAB's QR decomposition for a real $n \times (n + 1)$ matrix is about $3n^3$, which implies about 40 Megaflops for the QR decomposition of the Jacobian matrix for this example. This simple analysis illustrates the improved efficiency of the the approach proposed in this paper over previous homotopy approaches based on the optimal projection equations. Increased efficiency was also due to better selection of the initial system.

6. Conclusions

This paper has presented a new homotopy algorithm for the synthesis of H_2 optimal reduced-order models based on directly solving the optimal projection equations. The previous optimal projection equations based homotopy algorithms (Zigić *et al.* 1993a) are numerically robust but suffer from large dimensionality. The number of variables associated with this approach is of order nn_m. By parameterizing the reduced-order model, the gradient-based homotopy algorithms (Ge *et al.* 1993a,1993b) are more computationally efficient, but may cause numerical ill-conditioning. By using the results of Richter *et al.* (1993) to efficiently solve a pair of coupled Lyapunov equations, the effective number of variables associated with this approach is reduced to $n_m(m + l)$, which is identical to the dimension of the homotopy Jacobian inverted in the minimal parameterization approach of Ge *et al.* (1993a, 1993b). The examples of the previous section illustrated some of the features of the various algorithm options and some effects of the initialization schemes.

Appendix A.   Formulation of the Derivative and Correction Equations

Before deriving the derivative and correction equations (4.14)–(4.17), we state the following useful properties about the derivatives of the contragredient transformation of $(\hat{Q}, \hat{P})$.

Note that it follows from (4.9)-(4.11) that $\tau(\lambda)$ can be expressed as

$$\tau(\lambda) = \hat{Q}(\lambda)U^{\mathrm{T}}(\lambda)\Lambda^{\dagger}(\lambda)U(\lambda) = \hat{Q}(\lambda)U_1(\lambda)\Omega^{-1}(\lambda)U_1^{\mathrm{T}}(\lambda), \tag{A.1}$$

or

$$\tau(\lambda) = W(\lambda)\Lambda^{\dagger}(\lambda)W^{\mathrm{T}}(\lambda)\hat{P}(\lambda) = W_1(\lambda)\Omega^{-1}(\lambda)W_1^{\mathrm{T}}(\lambda)\hat{P}(\lambda), \tag{A.2}$$

where

$$\Lambda^{\dagger}(\lambda) = \begin{bmatrix} \Omega^{-1}(\lambda) & 0 \\ 0 & 0 \end{bmatrix}. \tag{A.3}$$

The representations of $\tau(\lambda)$ given by (A.1) and (A.2) are used below as a convenient way of expressing the derivative equations partially in terms of $\dot{\hat{Q}}(\lambda)$ and $\dot{\hat{P}}(\lambda)$ as opposed to expressing the derivative equations only in terms of $\dot{W}_1(\lambda), \dot{U}_1(\lambda)$, and $\dot{\Omega}(\lambda)$.

Differentiating (A.1) or (A.2), gives the following expressions for $\dot{\tau}$

$$\dot{\tau} = \dot{\hat{Q}}U_1\Omega^{-1}U_1^{\mathrm{T}} + \hat{Q}(\dot{U}_1\Omega^{-1}U_1^{\mathrm{T}} + U_1\Omega^{-1}\dot{\Omega}\Omega^{-1} + U_1\Omega^{-1}\dot{U}_1^{\mathrm{T}}), \tag{A.4}$$

or

$$\dot{\tau} = W_1\Omega^{-1}W_1^{\mathrm{T}}\dot{\hat{P}} + (W_1\Omega^{-1}\dot{W}_1^{\mathrm{T}} + W_1\Omega^{-1}\dot{\Omega}\Omega^{-1}W_1^{\mathrm{T}} + \dot{W}_1\Omega^{-1}W_1^{\mathrm{T}})\hat{P}, \tag{A.5}$$

with

$$\frac{d\Omega^{-1}}{d\lambda} = -[\Omega^{-1}]^2\dot{\Omega} = -\Omega^{-1}\dot{\Omega}\Omega^{-1}, \tag{A.6}$$

since Ω is diagonal. Below, we derive the matrix equations that can be used to solve for the derivatives and corrections.

A.1.   The Derivative Equations

Differentiating (4.5) and (4.6) and using (A.4)-(A.6), yields

$$0 = A_u\dot{\hat{Q}} + \dot{\hat{Q}}A_u^{\mathrm{T}} + R_{\hat{Q}} + R_{\hat{Q}}^{\mathrm{T}} + V_{\hat{Q}} + V_{\hat{Q}}^{\mathrm{T}}, \tag{A.7}$$

where

$$A_u \triangleq A + (I_n - \tau)\Sigma U_1\Omega^{-1}U_1^{\mathrm{T}}, \tag{A.8}$$

$$R_{\hat{Q}} \triangleq \hat{Q}(\dot{U}_1\Omega^{-1}U_1^{\mathrm{T}} - U_1\Omega^{-1}\dot{\Omega}\Omega^{-1}U_1^{\mathrm{T}} + U_1\Omega^{-1}\dot{U}_1^{\mathrm{T}})\Sigma(I_n - \tau)^{\mathrm{T}}, \tag{A.9}$$

$$V_{\hat{Q}} \triangleq [\dot{A} + (I_n - \tfrac{1}{2}\tau)\dot{\Sigma}U_1\Omega^{-1}U_1^{\mathrm{T}}]\hat{Q}, \tag{A.10}$$

22

and

$$0 = A_w^T \dot{P} + \dot{P} A_w + R_{\dot{P}} + R_{\dot{P}}^T + V_{\dot{P}} + V_{\dot{P}}^T, \tag{A.11}$$

where

$$A_w \triangleq A + W_1 \Omega^{-1} W_1^T \bar{\Sigma}(I_n - \tau), \tag{A.12}$$

$$R_{\dot{P}} \triangleq \hat{P}(\dot{W}_1 \Omega^{-1} W_1^T - W_1 \Omega^{-1} \dot{\Omega} \Omega^{-1} W_1^T + W_1 \Omega^{-1} \dot{W}_1^T)\bar{\Sigma}(I_n - \tau), \tag{A.13}$$

$$V_{\dot{P}} \triangleq \hat{P}[\dot{A} + W_1 \Omega^{-1} W_1^T \dot{\bar{\Sigma}}(I_n - \frac{1}{2}\tau)]. \tag{A.14}$$

Note that it follows from (4.1) that

$$\dot{A} = A_f - A_0, \quad \dot{B} = B_f - B_0, \quad \dot{C} = C_f - C_0,$$

$$\dot{V} = V_f - V_0, \quad \dot{R} = R_f - R_0,$$

$$\dot{\Sigma} = \dot{B}VB^T + B\dot{V}B^T + BV\dot{B}^T, \quad \dot{\bar{\Sigma}} = \dot{C}^T RC + C^T \dot{R}C + C^T R\dot{C}.$$

Next, differentiating (4.9) and (4.10), yields

$$\dot{Q} = \dot{W}_1 \Omega W_1^T + W_1 \dot{\Omega} W_1^T + W_1 \Omega \dot{W}_1^T, \tag{A.15}$$

$$\dot{P} = \dot{U}_1 \Omega U_1^T + U_1 \dot{\Omega} U_1^T + U_1 \Omega \dot{U}_1^T. \tag{A.16}$$

Furthermore, differentiating (4.12b) with respect to λ gives

$$0 = \dot{U}W + U\dot{W} = \dot{U}_1^T W_1 + U_1^T \dot{W}_1. \tag{A.17}$$

A.2. The Correction Equations

The correction equations are developed with λ at some fixed value. say λ^*. The derivation of the correction equation is based on the relationship between Newton's method and a particular homotopy. Below, we use the notation

$$\dot{f}(\theta) \triangleq \frac{\partial f}{\partial \theta}. \tag{A.18}$$

Let $f : \mathbf{R}^n \to \mathbf{R}^n$ be C^1 continuous and consider the equation

$$0 = f(\theta). \tag{A.19}$$

23

If $\theta^{(i)}$ is the current approximation to the solution of (A.19), then the Newton correction (Fletcher 1987) $\Delta\theta$ is given by

$$\theta^{(i+1)} - \theta^{(i)} \triangleq \Delta\theta = -\dot{f}(\theta^{(i)})^{-1}e, \qquad (A.20)$$

where

$$e \triangleq f(\theta^{(i)}). \qquad (A.21)$$

Now, let $\theta^{(i)}$ be an approximation to θ satisfying (A.19). Then with e given by (A.21) construct the following homotopy to solve (A.19)

$$(1-\beta)e = f(\theta(\beta)), \quad \beta \in [0,1]. \qquad (A.22)$$

Note that at $\beta = 0$, (A.22) has solution $\theta(0) = \theta^{(i)}$ while $\theta(1)$ satisfies (A.19). Then differentiating (A.22) with respect to β gives

$$\frac{\partial\theta}{\partial\beta}|_{\beta=0} = -\dot{f}(\theta^{(i)})^{-1}e. \qquad (A.23)$$

Remark A.1. Note that the Newton correction $\Delta\theta$ in (A.20) and the derivative $\frac{\partial\theta}{\partial\beta}|_{\beta=0}$ in (A.23) are identical. Hence, the Newton correction $\Delta\theta$ can be found by constructing a homotopy of the form (A.22) and solving for the resulting derivative $\frac{\partial\theta}{\partial\beta}|_{\beta=0}$. As seen below, this insight is particularly useful when deriving Newton corrections for equations that have a matrix structure.

Now, we use the insights of Remark A.1 to derive the equations that need to be solved for the Newton corrections $(\Delta\hat{Q}, \Delta\hat{P})$, or, equivalently, $(\Delta W_1, \Delta U_1, \Delta\Omega)$. We begin by recalling that λ is assumed to have some fixed value, say λ^*. Also, it is assumed that $(\hat{Q}^*, \hat{P}^*, W_1^*, U_1^*, \Omega^*)$ is the current approximation of $(\hat{Q}(\lambda^*), \hat{P}(\lambda^*), W_1(\lambda^*), U_1(\lambda^*), \Omega(\lambda^*))$ and that $E_{\hat{Q}}^*$ and $E_{\hat{P}}^*$ are respectively the errors in equations (4.5) and (4.6) with $\lambda = \lambda^*$ and $\hat{Q}(\lambda)$ and $\hat{P}(\lambda)$ replaced by $\hat{Q}^*$ and $\hat{P}^*$, respectively.

Next, we form the homotopy

$$(1-\beta)E_{\hat{Q}}^* = A\hat{Q}(\beta) + \hat{Q}(\beta)A^\mathrm{T} + \tau(\beta)\Sigma + \Sigma\tau^\mathrm{T}(\beta) + \tau(\beta)\Sigma\tau^\mathrm{T}(\beta). \qquad (A.24)$$

$$(1-\beta)E_{\hat{P}}^* = A^\mathrm{T}\hat{P}(\beta) + \hat{P}(\beta)A + \tau^\mathrm{T}(\beta)\bar{\Sigma} + \bar{\Sigma}\tau(\beta) + \tau^\mathrm{T}(\beta)\bar{\Sigma}\tau(\beta). \qquad (A.25)$$

Here, $(A, B, C, R, V) = (A(\lambda^*), B(\lambda^*), C(\lambda^*), R(\lambda^*), V(\lambda^*))$, i.e., the system matrices are assumed to be evaluated at $\lambda = \lambda^*$ and at $\beta = 0$, $(\hat{Q}(0), \hat{P}(0), \tau(0))$ is the current approximation. Differentiating (A.24) and (A.25) with respect to β, noting the identity of (A.4)-(A.6) with $\dot{\tau}$ now representing $\frac{d\tau}{d\beta}$, and using Remark A.1 to make the replacements

$$\Delta\hat{Q} \triangleq \frac{d\hat{Q}}{d\beta}|_{\beta=0}, \quad \Delta\hat{P} \triangleq \frac{d\hat{P}}{d\beta}|_{\beta=0}, \qquad (A.26a,b)$$

24

yields

$$0 = A_u \Delta \hat{Q} + \Delta \hat{Q} A_u^{\mathrm{T}} + R_{\Delta \hat{Q}} + R_{\Delta \hat{Q}}^{\mathrm{T}} + E_{\hat{Q}}^*, \qquad (A.27)$$

where

$$A_u \triangleq A + (I_n - \tau) \Sigma U_1 \Omega^{-1} U_1^{\mathrm{T}},$$

$$R_{\Delta \hat{Q}} \triangleq \hat{Q}(\Delta U_1 \Omega^{-1} U_1^{\mathrm{T}} - U_1 \Omega^{-1} \Delta \Omega \Omega^{-1} U_1^{\mathrm{T}} + U_1 \Omega^{-1} \Delta U_1^{\mathrm{T}}) \Sigma (I_n - \tau)^{\mathrm{T}},$$

and

$$0 = A_w^{\mathrm{T}} \Delta \hat{P} + \Delta \hat{P} A_w + R_{\Delta \hat{P}} + R_{\Delta \hat{P}}^{\mathrm{T}} + E_{\hat{P}}^*, \qquad (A.28)$$

where

$$A_w \triangleq A + W_1 \Omega^{-1} W_1^{\mathrm{T}} \bar{\Sigma} (I_n - \tau),$$

$$R_{\Delta \hat{P}} \triangleq \hat{P}(\Delta W_1 \Omega^{-1} W_1^{\mathrm{T}} - W_1 \Omega^{-1} \Delta \Omega \Omega^{-1} W_1^{\mathrm{T}} + W_1 \Omega^{-1} \Delta W_1^{\mathrm{T}}) \bar{\Sigma} (I_n - \tau).$$

Next, replacing λ with β in (4.9), (4.10), and (4.12) and differentiating them with respect to β, the following equations are derived,

$$\Delta \hat{Q} = \Delta W_1 \Omega W_1^{\mathrm{T}} + W_1 \Delta \Omega W_1^{\mathrm{T}} + W_1 \Omega \Delta W_1^{\mathrm{T}}, \qquad (A.29)$$

$$\Delta \hat{P} = \Delta U_1 \Omega U_1^{\mathrm{T}} + U_1 \Delta \Omega U_1^{\mathrm{T}} + U_1 \Omega \Delta U_1^{\mathrm{T}}, \qquad (A.30)$$

$$0 = \Delta U_1^{\mathrm{T}} W_1 + U_1^{\mathrm{T}} \Delta W_1, \qquad (A.31)$$

where

$$\Delta U_1 \triangleq \frac{dU}{d\beta}|_{\beta=0}, \quad \Delta W_1 \triangleq \frac{dW}{d\beta}|_{\beta=0}, \quad \Delta \Omega \triangleq \frac{d\Omega}{d\beta}|_{\beta=0}. \qquad (A.32)$$

Appendix B. Efficient Computation of the Solution to the Prediction and Correction Equations

This appendix presents a solution procedure using Richter *et al.* (1993) for efficiently solving the prediction equations (4.14)–(4.15) and the correction equations (4.16)–(4.17). We commence by recognizing that (4.14)–(4.15) and (4.16)–(4.17) have the following generic form:

$$0 = A_u \dot{Q} + \dot{Q} A_u^{\mathrm{T}} + \mathcal{F}_1(\dot{Q}, \dot{P}) + F_{\dot{Q}}, \qquad (B.1)$$

$$0 = A_w^{\mathrm{T}} \dot{P} + \dot{P} A_w + \mathcal{F}_2(\dot{Q}, \dot{P}) + F_{\dot{P}}. \qquad (B.2)$$

25

where the linear operators $\mathcal{F}_1 : X_1^2 S^{n\times n} \to S^{n\times n}$ and $\mathcal{F}_2 : X_1^2 S^{n\times n} \to S^{n\times n}$ are defined by

$$\mathcal{F}_1(\dot{Q}, \dot{P}) \triangleq R_{\dot{Q}}(\dot{Q}, \dot{P}) + R_{\dot{Q}}^T(\dot{Q}, \dot{P}), \qquad (B.3)$$

$$\mathcal{F}_2(\dot{Q}, \dot{P}) \triangleq R_{\dot{P}}(\dot{Q}, \dot{P}) + R_{\dot{P}}^T(\dot{Q}, \dot{P}), \qquad (B.4)$$

and $F_{\dot{Q}}$ and $F_{\dot{P}}$ are constant forcing terms. It is easy to verify that $(\dot{Q}, \dot{P}, \mathcal{F}_1, \mathcal{F}_2, F_{\dot{Q}}, F_{\dot{P}})$ in the above equations represents $(\dot{Q}, \dot{P}, R_{\dot{Q}} + R_{\dot{Q}}^T, R_{\dot{P}} + R_{\dot{P}}^T, V_{\dot{Q}} + V_{\dot{Q}}^T + E_{\dot{Q}_0}, V_{\dot{P}} + V_{\dot{P}}^T + E_{\dot{P}_0})$ in (4.14)–(4.15) and $(\Delta\dot{Q}, \Delta\dot{P}, R_{\Delta\dot{Q}} + R_{\Delta\dot{Q}}^T, R_{\Delta\dot{P}} + R_{\Delta\dot{P}}^T, E_{\dot{Q}}^*, E_{\dot{P}}^*)$ in (4.16)–(4.17), respectively. Our goal now is to find for some integers m_1 and m_2 (as small as possible) linear operators $\phi_1 : X_1^2 S^{n\times n} \to R^{m_1}$, $\mathcal{G}_1 : R^{m_1} \to S^{n\times n}$, $\phi_2 : X_1^2 S^{n\times n} \to R^{m_2}$, $\mathcal{G}_2 : R^{m_2} \to S^{n\times n}$, such that

$$\mathcal{F}_1(\dot{Q}, \dot{P}) = \mathcal{G}_1(\phi_1(\dot{Q}, \dot{P})), \qquad (B.5a)$$

$$\mathcal{F}_2(\dot{Q}, \dot{P}) = \mathcal{G}_2(\phi_2(\dot{Q}, \dot{P})). \qquad (B.5b)$$

First, let T_u and T_w be the transformation matrices such that $T_u^{-1} A_u T_u$ and $T_w^{-1} A_w T_w$ are in suitable form according to the *basis* option described in Section 4.3.2. Next, make the replacements

$$A_u \leftarrow T_u^{-1} A_u T_u, \quad A_w \leftarrow T_w^{-1} A_w T_w,$$

$$\dot{Q} \leftarrow T_u^{-1} \dot{Q} T_u^{-T}, \quad \dot{P} \leftarrow T_w^T \dot{P} T_w,$$

$$F_{\dot{Q}} \leftarrow T_u^{-1} F_{\dot{Q}} T_u^{-T}, \quad F_{\dot{P}} \leftarrow T_w^T F_{\dot{P}} T_w,$$

$$W_1 \leftarrow T_u^{-1} W_1, \quad U_1 \leftarrow T_w^T U_1,$$

$$W_2 \leftarrow T_u^{-1} W_2, \quad U_2 \leftarrow T_w^T U_2,$$

$$B \leftarrow T_w^{-1} B, \quad C \leftarrow C T_u.$$

Then, in this new basis we obtain

$$R_{\dot{Q}}(\dot{Q}, \dot{P}) = W_1 \Omega W_1^T T_u^T T_w^{-T} (\dot{U}_1 \Omega^{-1} U_1^T - U_1 \Omega^{-1} \dot{\Omega} \Omega^{-1} U_1^T + U_1 \Omega^{-1} \dot{U}_1^T) BVB^T U_2 W_2^T, \quad (B.6a)$$

and

$$R_{\dot{P}}(\dot{Q}, \dot{P}) = U_1 \Omega U_1^T T_w^{-1} T_u (\dot{W}_1 \Omega^{-1} W_1^T - W_1 \Omega^{-1} \dot{\Omega} \Omega^{-1} W_1^T + W_1 \Omega^{-1} \dot{W}_1^T W_1^T) C^T R C W_2 U_2^T. \quad (B.6b)$$

Now, rewrite (B.6a) as

$$R_{\dot{Q}}(\dot{Q}, \dot{P}) = F_A S_L (\dot{U}_1 \Omega^{-1} U_1^T - U_1 \Omega^{-1} \dot{\Omega} \Omega^{-1} U_1^T + U_1 \Omega^{-1} \dot{U}_1^T) S_R G_A, \qquad (B.7)$$

26

where for some integers $m_s < n$ and $n_s < n$

$$F_A \in \mathbf{R}^{n \times m_s}, \quad S_L \in \mathbf{R}^{m_s \times n}, \quad S_R \in \mathbf{R}^{n \times n_s}, \quad G_A \in \mathbf{R}^{n_s \times n}. \qquad (B.7a)$$

In a similar way, we can express $R_{\dot{P}}(\dot{\hat{Q}}, \dot{P})$ as

$$R_{\dot{P}}(\dot{\hat{Q}}, \dot{P}) = F_B T_L (\dot{W}_1 \Omega^{-1} W_1^T - W_1 \Omega^{-1} \dot{\Omega} \Omega^{-1} W_1^T + W_1 \Omega^{-1} \dot{W}_1^T W_1^T) T_R G_B, \qquad (B.8)$$

where for some integers $m_t < n$ and $n_t < n$

$$F_B \in \mathbf{R}^{n \times m_t}, \quad T_L \in \mathbf{R}^{m_t \times n}, \quad T_R \in \mathbf{R}^{n \times n_t}, \quad G_B \in \mathbf{R}^{n_t \times n}. \qquad (B.8a)$$

The choice of (F_A, S_L, S_R, G_A) in (B.7) and (F_B, T_L, T_R, G_B) in (B.8) are not unique. The solution procedure we discuss below is most efficient if we minimize the products $m_s n_s$ and $m_t n_t$.

Using (B.7), it follows that $\phi_1(\cdot)$ and $\mathcal{G}_1(\cdot)$ in (B.5) can be defined by

$$\phi_1(\dot{\hat{Q}}, \dot{P}) \triangleq \mathrm{vec}(S_L [\dot{U}_1 \Omega^{-1} U_1^T - U_1 \Omega^{-1} \dot{\Omega} \Omega^{-1} U_1^T + U_1 \Omega^{-1} \dot{U}_1^T] S_R), \qquad (B.9)$$

and

$$\mathcal{G}_1(z) \triangleq F_A \, \mathrm{vec}^{-1}(z) G_A + G_A^T [\, \mathrm{vec}^{-1}(z)]^T F_A^T, \qquad (B.10)$$

such that

$$m_1 = m_s n_s. \qquad (B.11)$$

Similarly, it follows from (B.8) that $\phi_2(\cdot)$ and $\mathcal{G}_2(\cdot)$ in (B.6) can be defined by

$$\phi_2(\dot{\hat{Q}}, \dot{P}) \triangleq \mathrm{vec}(T_L [\dot{W}_1 \Omega^{-1} W_1^T - W_1 \Omega^{-1} \dot{\Omega} \Omega^{-1} W_1^T + W_1 \Omega^{-1} \dot{W}_1^T] T_R), \qquad (B.12)$$

and

$$\mathcal{G}_2(z) \triangleq F_B \, \mathrm{vec}^{-1}(z) G_B + G_B^T [\, \mathrm{vec}^{-1}(z)]^T F_B^T, \qquad (B.13)$$

such that

$$m_2 = m_t n_t. \qquad (B.14)$$

Note that it is assumed in (B.9) and (B.12) that $\dot{W}_1, \dot{U}_1$, and $\dot{\Omega}$ are obtained from (A.15)–(A.17), or, equivalently, (A.29)–(A.31). A procedure to compute $\dot{W}_1, \dot{U}_1$, and $\dot{\Omega}$ given $\dot{\hat{Q}}$ and $\dot{P}$ is presented in Appendix C.

Now, with the definitions (B.9), (B.10), (B.12), and (B.13), the solution procedure for coupled, modified Lyapunov equations described in Richter *et al.* (1993) is applied to solve for $(\dot{\hat{Q}}, \dot{P})$ in (B.1)

27

and (B.2). With the above formulation, the efficiency of the coupled Lyapunov equations solver is realized Richter *et al.* (1993) by exploiting the low rank properties of the coupling terms. Here, we illustrate this by an example solving the prediction equations. Suppose $m, l < (n - n_m) < n$, i.e., the number of the inputs and outputs are less than the difference between the order of the original plant and the desired reduced-order. Noting that (B.7) and (B.8) are equivalent to (A.9) and (A.13), respectively, to minimize $m_1 = m_s n_s$ and $m_2 = m_t n_t$, we choose

$$F_A = W_1 \Omega, \quad S_L = W_1^T T_u^T T_w^{-T}, \quad S_R = B, \quad G_A = V B^T U_2 W_2^T,$$

in (B.7), and

$$F_B = U_1 \Omega, \quad T_L = U_1^T T_w^{-1} T_u, \quad T_R = C^T, \quad G_B = RC(I_n - \tau),$$

in (B.8). Thus, using (B.7a) and (B.8a), it follows that $m_s = n_m$, $n_s = l$, $m_t = n_m$, $n_t = m$, which results in $m_* = m_1 + m_2 = n_m(m + l)$. Now, using the solution procedure described in Richter *et al.* (1993), to solve the prediction equations (B.1) and (B.2), the primary computation burden is to invert a matrix of dimension $m_* \times m_*$ and to solve two sets of $m_* + 1$ standard $n \times n$ Lyapunov equations, with one set having A_u as the coefficient matrix and the other set with coefficient matrix A_w.

In comparison, by using Kronecker algebra (Brewer 1978) and exploiting the rank condition (2.18) of $\hat{Q}$ and $\hat{P}$, (B.1) and (B.2) can be converted to the vector form $\mathcal{A}\chi = b$ where χ is a vector consisting of the independent elements of W_1, U_1 and Ω given by (4.12) and (4.13). Hence, to get the solution for $(\dot{\hat{Q}}, \dot{\hat{P}})$, it is required to invert an $(2nn_m + n_m^2) \times (2nn_m + n_m^2)$ matrix. The approach proposed in Zigić *et al.* (1993a) involves the computation of the kernel of a Jacobian matrix. The Jacobian matrix has $2nn_m + n_m^2$ rows and $2nn_m + n_m^2 + 1$ columns. The kernel is found by computing a QR factorization of the Jacobian matrix then using a back substitution. Thus, if $m << n$ and $l << n$, which is usually true, m_* is sufficiently small and the algorithm discussed in this Appendix will be much more efficient. Furthermore, if (B.1) and (B.2) are first transformed to the bases in which A_u and A_w are nearly diagonal, respectively, the cost of computation can be reduced significantly. The comparison in computation time is discussed in Example 5.3.

28

Appendix C.   Conversion from $(\dot{\underline{Q}}, \dot{\underline{P}})$ to $(\dot{W}_1, \dot{U}_1, \dot{\Omega})$

Note that the following procedure is valid only in the original basis. It is desired to compute $\dot{W}_1$, $\dot{U}_1$ and $\dot{\Omega}$ satisfying (A.15)-(A.17). Note that (A.15) implies

$$U_1^T W_1 = I_{n_m}, \quad U_1^T W_2 = 0, \quad U_2^T W_1 = 0. \tag{C.1}$$

Pre- and post-multiplying (4.22) by U and U^T respectively gives

$$\dot{\underline{Q}} = \underline{\dot{W}}_1 [\Omega \quad 0] + \begin{bmatrix} \dot{\Omega} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Omega \\ 0 \end{bmatrix} (\underline{\dot{W}}_1)^T, \tag{C.2}$$

where

$$\dot{\underline{Q}} \triangleq U \dot{Q} U^T, \tag{C.3}$$

$$\underline{\dot{W}}_1 \triangleq U \dot{W}_1 = \begin{bmatrix} U_1^T \dot{W}_1 \\ U_2^T \dot{W}_2 \end{bmatrix}. \tag{C.4}$$

Similarly, pre- and post-multiplying (A.16) by W^T and W respectively gives

$$\dot{\underline{P}} = \underline{\dot{U}}_1 [\Omega \quad 0] + \begin{bmatrix} \dot{\Omega} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Omega \\ 0 \end{bmatrix} \underline{\dot{U}}_1^T, \tag{C.5}$$

where

$$\dot{\underline{P}} \triangleq W^T \dot{P} W, \tag{C.6}$$

$$\underline{\dot{U}}_1 \triangleq W^T \dot{U}_1 = \begin{bmatrix} W_1^T \dot{U}_1 \\ W_2^T \dot{U}_1 \end{bmatrix}. \tag{C.7}$$

Partition $\underline{\dot{W}}_1$ and $\underline{\dot{U}}_1$ as

$$\underline{\dot{W}}_1 = \begin{bmatrix} \underline{\dot{W}}_{11} \\ \underline{\dot{W}}_{21} \end{bmatrix}, \tag{C.8}$$

$$\underline{\dot{U}}_1 = \begin{bmatrix} \underline{\dot{U}}_{11} \\ \underline{\dot{U}}_{21} \end{bmatrix}. \tag{C.9}$$

It then follows from (C.4) and (C.7)-(C.9) that (A.17) is equivalent to

$$\underline{\dot{W}}_{11} = -\underline{\dot{U}}_{11}^T. \tag{C.10}$$

29

It now follows from (C.8) that (C.2) is equivalent to

$$
\begin{bmatrix} \dot{\underline{Q}}_{11} & \dot{\underline{Q}}_{21})^{\mathrm{T}} \\ \dot{\underline{Q}}_{21} & 0 \end{bmatrix} = \begin{bmatrix} \dot{\underline{W}}_{11}\Omega & 0 \\ \dot{\underline{W}}_{21}\Omega & 0 \end{bmatrix} + \begin{bmatrix} \dot{\Omega} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Omega(\dot{\underline{W}}_{11})^{\mathrm{T}} & \Omega(\dot{\underline{W}}_{21})^{\mathrm{T}} \\ 0 & 0 \end{bmatrix},
\qquad (C.11)
$$

and from (C.9) that (C.5) is equivalent to

$$
\begin{bmatrix} \dot{\underline{P}}_{11} & (\dot{\underline{P}}_{21})^{\mathrm{T}} \\ \dot{\underline{P}}_{21} & 0 \end{bmatrix} = \begin{bmatrix} \dot{\underline{U}}_{11}\Omega & 0 \\ \dot{\underline{U}}_{21}\Omega & 0 \end{bmatrix} + \begin{bmatrix} \dot{\Omega} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Omega(\dot{\underline{U}}_{11})^{\mathrm{T}} & \Omega(\dot{\underline{U}}_{21})^{\mathrm{T}} \\ 0 & 0 \end{bmatrix}.
\qquad (C.12)
$$

Furthermore, (C.11) is equivalent to

$$
\dot{\underline{Q}}_{11} = \dot{\underline{W}}_{11}\Omega + \dot{\Omega} + \Omega(\dot{\underline{W}}_{11})^{\mathrm{T}},
\qquad (C.13)
$$

$$
\dot{\underline{Q}}_{21} = \dot{\underline{W}}_{21}\Omega.
\qquad (C.14)
$$

Similarly, equation (C.12) is equivalent to

$$
\dot{\underline{P}}_{11} = \dot{\underline{U}}_{11}\Omega + \dot{\Omega} + \Omega(\dot{\underline{U}}_{11})^{\mathrm{T}},
\qquad (C.15)
$$

$$
\dot{\underline{P}}_{21} = \dot{\underline{U}}_{21}\Omega.
\qquad (C.16)
$$

Now, (C.14) and (C.16) imply respectively that

$$
\dot{\underline{W}}_{21} = \dot{\underline{Q}}_{21}\Omega^{-1},
\qquad (C.17)
$$

$$
\dot{\underline{U}}_{21} = \dot{\underline{P}}_{21}\Omega^{-1}.
\qquad (C.18)
$$

Furthermore, substituting (C.10) into (C.13) yields

$$
\dot{\underline{Q}}_{11} = -(\dot{\underline{U}}_{11})^{\mathrm{T}}\Omega + \dot{\Omega} - \Omega\dot{\underline{U}}_{11}.
\qquad (C.19)
$$

Denote the (i, j) elements of $\dot{\underline{P}}_{11}$, $\dot{\underline{Q}}_{11}$, and $\dot{\underline{U}}_{11}$ respectively by $\dot{\underline{p}}_{ij}$, $\dot{\underline{q}}_{ij}$, and $\dot{\underline{u}}_{ij}$. Then we can rewrite (C.15) and (C.19) as

$$
\dot{\underline{p}}_{ij} = \dot{\underline{u}}_{ij}\omega_j + \delta_{ij}\dot{\omega}_i + \omega_i\dot{\underline{u}}_{ji}, \quad i, j \in \{1, 2, \ldots, n_m\},
\qquad (C.20)
$$

$$
\dot{\underline{q}}_{ij} = -\dot{\underline{u}}_{ji}\omega_j + \delta_{ij}\dot{\omega}_i - \omega_i\dot{\underline{u}}_{ij}, \quad i, j \in \{1, 2, \ldots, n_m\},
\qquad (C.21)
$$

where

$$
\delta_{ij} \triangleq \begin{cases} 1 \text{ for } i = j \\ 0 \text{ for } i \neq j. \end{cases}
\qquad (C.22)
$$

30

Next, assume $i = j$. Then subtracting (C.21) from (C.20) gives

$$\underline{\dot{u}}_{ii} = \frac{\underline{\dot{p}}_{ii} - \underline{\dot{q}}_{ii}}{4\omega_i}. \qquad (C.23)$$

Now, assume $i \neq j$. Multiplying (C.20) by (ω_j/ω_i) and adding the resultant equations to (C.21) gives

$$\underline{\dot{u}}_{ij} = \frac{\omega_j \underline{\dot{p}}_{ij} + \omega_i \underline{\dot{q}}_{ij}}{\omega_j^2 - \omega_i^2}, \quad \omega_i \neq \omega_j, \qquad (C.24a)$$

or, if $\omega_i = \omega_j$,

$$\underline{\dot{u}}_{ij} = \frac{\underline{\dot{p}}_{ij} - \underline{\dot{q}}_{ji}}{4\omega_i}. \qquad (C.24b)$$

Now, $\underline{\dot{U}}_{11}$ is defined by (C.23) and (C.24) and $\underline{\dot{U}}_{21}$ by (C.18). $\underline{\dot{W}}_{11}$ is then defined by (C.10) and $\underline{\dot{W}}_{21}$ by (C.17). $\underline{\dot{W}}_1$ and $\underline{\dot{U}}_1$ are now defined respectively by (C.8) and (C.9). Using (A.17) it follows from (C.4) and (C.7) that $\dot{W}_1$ and $\dot{U}_1$ are given respectively by

$$\dot{W}_1 = W \underline{\dot{W}}_1, \qquad (C.25)$$

$$\dot{U}_1 = U^{\mathrm{T}} \underline{\dot{U}}_1. \qquad (C.26)$$

From (C.22) it follows that

$$\dot{\omega}_i = \frac{\underline{\dot{p}}_{ii} + \underline{\dot{q}}_{ii}}{2}, \qquad (C.27)$$
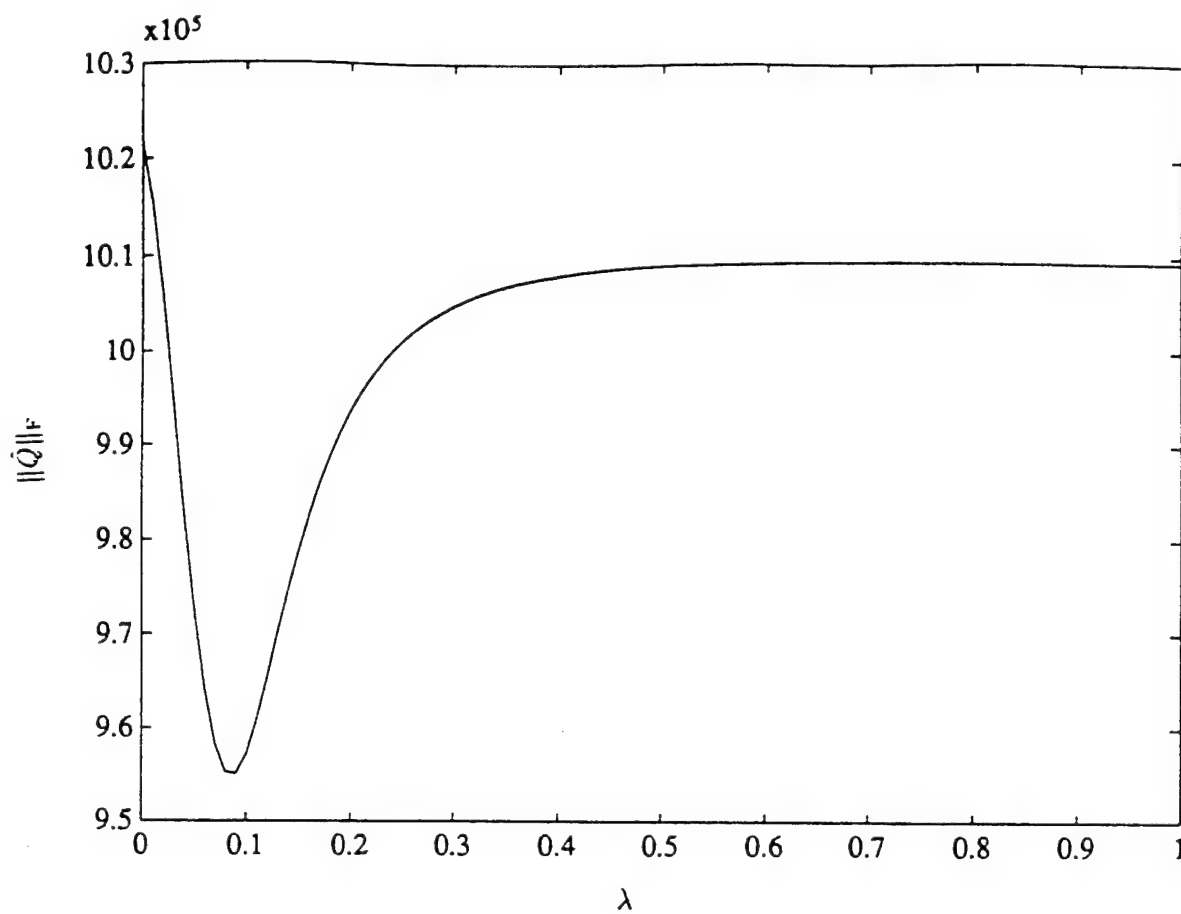
which defines $\dot{\Omega}$.

31
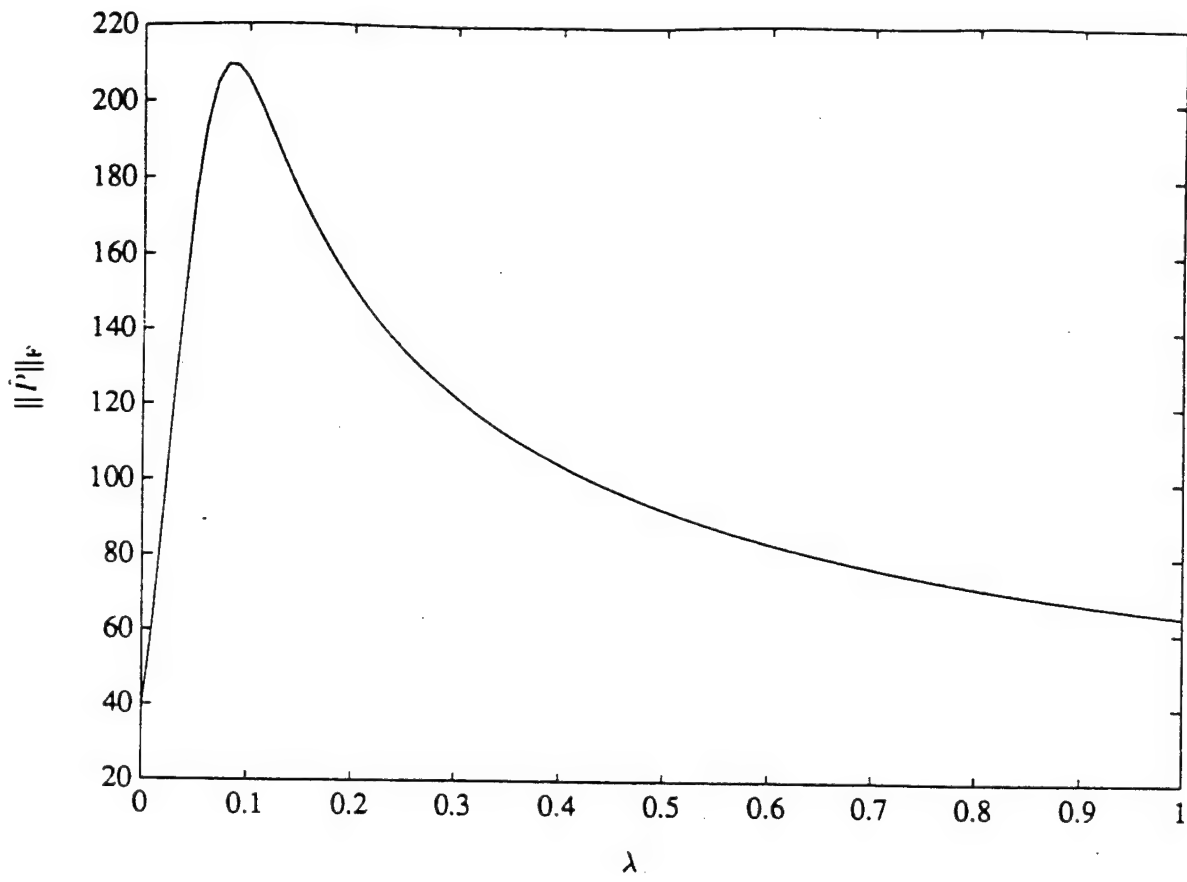
Figure 5.1. $\|\hat{Q}\|_F$ vs λ for Example 5.2

Figure 5.2. $\|\hat{P}\|_F$ vs λ for Example 5.2

References

Alexander, J. C., and Yorke, J. A., 1978, The homotopy continuation method: numerically implementable topological procedures. *Transactions of the American Mathematical Society,* **242**, 271–284.

Aplevich, J.D., 1973, Gradient methods for optimal linear system reduction. *Int. J. Control,* **18**, 767–772.

Avila, J. H., 1974, The feasibility of continuation methods for nonlinear equations. *SIAM Journal of Numerical Analysis,* **11**, 104–144.

Bernstein, D.S. and Haddad, W.M., 1990, Robust stability and performance via fixed-order dynamic compensation with guaranteed cost bounds. *Math. Control Signal Systems,* **3**, 139–163.

Brewer, J.W., 1978, Kronecker products and matrix calculus in system theory. *IEEE Trans. Circuit and Systems,* **25**, 772–781.

Bryson, A.E., Jr. and Carrier, A., 1990, Second-order algorithm for optimal model order reduction. *Journal of Guidance, Control and Dynamics,* 887–892.

Collins, E. G., Jr., Davis, L.D., and Richter, S., 1993, A homotopy algorithm for maximum entropy design. *Proc. Amer. Contr. Conf.,* San Francisco, CA, 1010–1019.

Collins, E. G., Jr., Phillips, D. J., and Hyland, D. C., 1991, Robust decentralized control laws for the ACES Structure. *Control Systems Magazine,* 62–70.

De Villemagne, C., and Skelton, R.E., 1987, Model reduction using a projection formulation. *Int. J. Control,* **46**, 2141–2169.

Wacker, H., 1978, *Continuation Methods.* Academic Press, New York.

Dunyak, J. P., Junkins, J. L., and Watson, L. T., 1984, Robust nonlinear least squares estimation using the Chow-Yorke homotopy method. *Journal of Guidance, Control and Dynamics,* **7**, 752–755.

Eaves, B. C., Gould, F. J., Peoitgen, J. O., and Todd, M. J., 1983, *Homotopy Methods and Global Convergence,* Plenum Press, New York.

Fletcher, R., 1987, *Practical Methods of Optimization: Second Edition.* John Wiley and Sons, New York.

Garcia, C. B., and Zangwill, W. I., 1981, *Pathways to Solutions, Fixed Points and Equilibria,* Prentice-Hall, Englewood Cliffs, NJ.

Ge, Y., Collins, E. G., Jr., Watson, L. T., and Davis, L. D., 1993a, A input normal form homotopy for the L^2 optimal model order reduction problem. submitted to *Int. J. Control.*

Ge, Y., Collins, E. G., Jr., Watson, L. T., and Davis, L. D., 1993b, A comparison of homotopies for alternative formulations of the L^2 optimal model order reduction problem. submitted to *J. Comp. Appl. Math.*

Hickin, J., and Sinha, N.K., 1980, Model reduction for linear multivariable systems. *IEEE*

decentralized state feedback. *IEEE Transactions on Automatic Control,* **29**, No. 2, 148–158.

Riggs, J. B., and Edgar, T. F., 1974, Least squares reduction of linear systems using impulse response. *Int. J. Control,* **20**, 213–223.

Sebok, D. R., Richter, S. and DeCarlo, R., 1986, Feedback gain optimization in decentralized eigenvalue assignment. *Automatica,* **22**, 433-447.

Shin, Y. S., Haftka, R. T., Watson, L. T., and Plaut, R. H., 1988, Tracking structural optima as a function of available resources by a homotopy method. *Computer Methods in Applied Mechanics and Engineering,* **70**, 151–164.

Skelton, K.E., and Kabamba, P., 1986, Comments on "balanced gains and their significance for L^2 model reduction". *IEEE Trans. Automat. Contr.,* **31**, 796–797.

Spanos, J.T., Milman, M.H., and Mingori, D.L., 1990, Optimal model reduction and frequency-weighted extension. *Journal of Guidance, Control and Dynamics,* 271–284.

Turner, J. D., and Chun, H. M., 1984, Optimal distributed control of a flexible spacecraft during a large-angle maneuver. *Journal of Guidance, Control and Dynamics,* **7**, 257–264.

Watson, L. T., 1986, Numerical linear algebra aspects of globally convergent homotopy methods. *SIAM Review,* **28**, 529–545.

Watson, L. T., 1987, ALGORITHM 652 HOMPACK: A suite of codes for globally convergent homotopy algorithms. *ACM Transactions on Mathematical Software,* **13**, 281–310.

Wilson, D.A., 1970, Optimum solution of model-reduction problem. *Proc. IEE,* **117**, 1161–1165.

Wilson, D.A., 1974, Model reduction for multivariable systems. *Int. J. Control,* **20**, 57–64.

Wilson, D.A., and Mishra, R.N., 1979, Optimal reduction of multivariable systems, *Int. J. Control,* **29**, 267–278.

Zigić, D., Watson, L. T., Collins, E. G., Jr., and Bernstein, D. S., 1992, Homotopy methods for solving the optimal projection equations for the H_2 reduced order model problem. *International Journal of Control,* **56**, 173–191.

Zigić, D., Watson, L. T., Collins, E. G., Jr., and Bernstein, D. S., 1993a, Homotopy approaches to the H_2 reduced order model problem. *Journal of Mathematical Systems, Estimation, Control.* to appear.

Zigić, D., Watson, L. T., and Collins, E. G., Jr., 1993b, A homotopy method for solving Riccati equations on a shared memory parallel computer. *Sixth SIAM Conference on Parallel Processing for Scientific Computing .* 614–617.

Appendix K:

Reduced-Order Dynamic Compensation Using
the Hyland and Bernstein Optimal Projection Equations

Reduced-Order Dynamic Compensation Using the Hyland-Bernstein Optimal Projection Equations

by

Emmanuel G. Collins, Jr.
Department of Mechanical Engineering
Florida A& M/Florida State
Tallahassee, FL 32316
(904) 487-6331
FAX: (904) 487-6337
ecollins@evax.eng.fsu.edu

Wassim M. Haddad
School of Aerospace Engineering
Georgia Institute of Technology
Atlanta, GA 30332-0150
(404) 894-1078
FAX: (404) 894-2760
wm.haddad@aerospace.gatech.edu

Sidney S. Ying
Rockwell International
Collins Commercial Avionics
MS 306-100
Melbourne, FL 32934
(407) 768-7063
FAX: (407) 254-7805
ssy@dllws.cca.cr.rockwell.com

Abstract

Gradient-based homotopy algorithms have previously been developed for synthesizing H_2 optimal reduced-order dynamic compensators. These algorithms are made efficient and avoid high-order singularities along the homotopy path by constraining the controller realization to a minimal parameter basis. However, the resultant homotopy algorithms sometimes experience numerical ill-conditioning or failure due to the minimal parameterization constraint. This paper presents a new homotopy algorithm which is based on solving the optimal projection equations, a set of coupled Riccati and Lyapunov equations that characterize the optimal reduced-order dynamic compensator. Path following in the proposed algorithm is accomplished using a predictor/corrector scheme that computes the prediction and correction steps by efficiently solving a set of four Lyapunov equations coupled by relatively low rank linear operators. The algorithm does not suffer from ill-conditioning due to constraining the controller basis and often exhibits better numerical properties than the gradient-based homotopy algorithms.

1. Introduction

The design of reduced-order dynamic compensators is of practical importance due to limitations on the throughput of control processors. Hence, an important research area has involved the development of techniques for synthesizing H_2 optimal reduced-order compensators. Most of the techniques for designing optimal reduced-order compensators have been gradient-based parameter optimization methods which represent the controller by some parameter vector and attempt to find a vector for which the gradient of the performance index is zero, or, equivalently, the cost functional is minimal.

In the survey paper by Makila and Toivonen[1], several gradient-based approaches were discussed. Levine-Athans-type algorithms[2-7] are based on using some standard optimization methods (e.g., conjugate gradient algorithms) to iteratively solve the necessary conditions of optimality which minimize the cost increment. This approach requires the solution of a nonlinear matrix equation at each correction step but guarantees a cost descent direction without a line search. The Anderson-Moore algorithm[8] is based on minimizing a quadratic, positive-definite approximation of the second-order Taylor series expansion of the cost function increment. The descent Anderson-Moore approach utilizes gradient search schemes to guarantee the cost is reduced at each iteration and enhance convergence to a stationary point of the cost function[9,10]. For Newton-like approaches[11], instead of approximating the Hessian of the cost functional with a positive-definite matrix, the *actual* second-order expansion is minimized which involves computing the Newton correction step as the solution of a system of linear matrix equations at each iteration.

Recently, homotopy algorithms have been developed for the synthesis of optimal reduced-order compensators[12-15]. A gradient-based algorithm has been developed[15] that is made efficient and avoids high order singularities along the homotopy path by constraining the controller realization to a minimal parameterization basis. These algorithms[15] sometimes exhibits numerical ill-conditioning or can even fail due to the basis constraint. This is because minimal parameterizations of a given form may not exist at each point along the homotopy path or may force the algorithm to be ill-conditioned when the transformation to the given basis is ill-conditioned. Nonminimal parameterizations exhibit singularities along the homotopy path that can be handled heuristically but may also lead to ill-conditioning. This ill-conditioning is also observed outside of the context of homotopy algorithms by Kuhn and Schmidt[16]. Similar conclusions are presented in Refs. 17 and 18 for the closely related H_2 optimal model reduction problem.

The homotopy algorithm of Ref. 19 was based on solving the optimal projection equations developed by Hyland and Bernstein[20]. The optimal projection equations are a set of coupled Riccati and Lyapunov equations that characterize optimal reduced-order dynamic compensators. The equations decouple and the Riccati equations specialize to the standard LQG Riccati equations when the compensator is constrained to be full-order. The initial homotopy algorithm[19] for solving the optimal projection equations utilized a very crude path following scheme in which the Riccati equations and Lyapunov equations were *not* updated simultaneously. This caused the algorithm to exhibit poor convergence properties, especially as the control authority was increased.

This paper presents a homotopy algorithm to solve the optimal projection equations that simultaneously updates the coupled Riccati and Lyapunov equations. The path following is accomplished using a predictor/corrector integration scheme that computes the prediction and correction steps by solving a set of four Lyapunov equations coupled by relatively low rank linear operators. These equations are solved efficiently by using the technique presented in Ref. 21. This helps to avoid the very large dimensionality of similar algorithms based on the optimal projection equations for H_2 model reduction[22,23]. A model reduction algorithm that uses a similar approach to that used here is found in Ref. 24. Also, a related algorithm for full-order Maximum Entropy robust design is presented in Ref. 25. These results all show that algorithms based on the optimal projection equations tend to avoid the numerical ill-conditioning experienced in gradient-based algorithms due to constraints on the realization of the reduced-order model or controller.

The current homotopy algorithm, unlike some of the previous algorithms[17,18,22,23], assumes that the homotopy curve is monotonic with respect to the homotopy parameter. As discussed in Ref. 13, this assumption may not always be satisfied. It appears to be possible to extend the algorithm to relax this assumption without significantly increasing the required computations by using a technique related to that developed in Ref. 26. However, this is a subject of future research.

The paper is organized as follows. Section 2 presents the optimal projection equations for the H_2 reduced-order control problem. Section 3 gives a brief synopsis of homotopy methods. Next, Section 4 develops a new homotopy algorithm for optimal reduced-order controller design based on the optimal projection equations. Section 5 illustrates the algorithm with two illustrative examples. Finally, Section 6 presents the conclusion.

Nomenclature

$\mathbb{E}$	expected value
$\mathbf{R}^n, \mathbf{R}^{m \times n}$	$n \times 1$ real vectors, $m \times n$ real matrices
$Y \geq X$	$Y - X$ is nonnegative definite
$Y > X$	$Y - X$ is positive definite
x_{ij} or $X_{i,j}$	(i,j) element of matrix X
$X^\dagger$	*Moore-Penrose generalized inverse*[27] of matrix X
$X^\#$	*Group inverse*[27] of matrix X
I_r	$r \times r$ identity matrix
tr X	trace of square matrix X
$\|X\|_F^2, \ \|X\|_A$	Frobenius norm $(\|X\|_F^2 \triangleq \text{tr } XX^T)$, absolute norm $(\|X\|_A \triangleq \max_{i,j}\|X_{i,j}\|)$
$\text{vec}(\cdot)$	the invertible linear operator defined such that
	$\text{vec}(S) \triangleq [s_1^T \ s_2^T \cdots s_q^T]^T, \ S \in \mathbf{R}^{p \times q}$,
	where $s_j \in \mathbf{R}^p$ denotes the jth column of S.
$e_m^{(i)}$	the m-dimensional column vector whose i^{th} element
	equals one and whose additional elements are zeros.
$X(:,k)$	k^{th} column of the matrix X ⟨MATLAB notation⟩
$\mathbf{S}^{n \times n}$	the space of symmetric matrices in $\mathbf{R}^{n \times n}$
$\mathbf{X}_1^2 \mathbf{S}^{n \times n}$	$\mathbf{S}^{n \times n} \times \mathbf{S}^{n \times n}$

2. H_2 Optimal Reduced-Order Dynamic Compensation

Consider the nth-order linear time-invariant plant

$$\dot{x}(t) = Ax(t) + Bu(t) + D_1 w(t), \tag{2.1}$$

$$y(t) = Cx(t) + Du(t) + D_2 w(t), \tag{2.2}$$

3

where (A, B) is stabilizable, (A, C) is detectable, $x \in \mathbf{R}^n, u \in \mathbf{R}^m, y \in \mathbf{R}^l$, and $w \in \mathbf{R}^d$ is a standard white noise disturbance with intensity I_d and rank $D_2 = l$. The intensities of $D_1 w(t)$ and $D_2 w(t)$ are thus given, respectively, by $V_1 \triangleq D_1 D_1^{\mathrm{T}} \geq 0$, and $V_2 \triangleq D_2 D_2^{\mathrm{T}} > 0$. For convenience, we assume that $V_{12} \triangleq D_1 D_2^{\mathrm{T}} = 0$, i.e., the plant disturbance and measurement noise are uncorrelated. The goal of the *optimal reduced-order dynamic compensation problem* is to determine an n_c^{th}-order dynamic compensator

$$\dot{x}_c(t) = A_c x_c(t) + B_c y(t), \qquad (2.3)$$

$$u(t) = -C_c x_c(t), \qquad (2.4)$$

which satisfies the following two design criteria:

(i) the closed-loop system corresponding to (2.1)–(2.4) given by

$$\dot{\tilde{x}}(t) = \tilde{A} \tilde{x}(t) + \tilde{D} w(t), \qquad (2.5)$$

where

$$\tilde{x}(t) \triangleq \begin{bmatrix} x(t) \\ x_c(t) \end{bmatrix}, \quad \tilde{A} \triangleq \begin{bmatrix} A & -BC_c \\ B_c C & A_c - B_c D C_c \end{bmatrix}, \quad \tilde{D} \triangleq \begin{bmatrix} D_1 \\ B_c D_2 \end{bmatrix}, \qquad (2.6)$$

is asymptotically stable; and

(ii) the steady-state quadratic performance criterion

$$J(A_c, B_c, C_c) \triangleq \lim_{t \to \infty} \frac{1}{t} \mathrm{I\!E} \int_0^t [x^{\mathrm{T}}(s) R_1 x(s) + u^{\mathrm{T}}(s) R_2 u(s)] ds, \qquad (2.7)$$

where $R_1 \geq 0$ and $R_2 > 0$, is minimized.

Although a cross-weighting term of the form $2 x^{\mathrm{T}}(t) R_{12} u(t)$ can also be included in (2.7), we shall not do so here to facilitate the presentation. With the first criterion, we restrict our attention to the set of stabilizing compensators, $\mathcal{S}_c \triangleq \{(A_c, B_c, C_c) : \tilde{A} \text{ is asymptotically stable}\}$ which guarantees that the cost J is finite and independent of initial conditions. The cost (2.7) can now be expressed as

$$J(A_c, B_c, C_c) = \lim_{t \to \infty} \mathrm{I\!E}[\tilde{x}^{\mathrm{T}}(t) \tilde{R} \tilde{x}(t)], \qquad (2.8)$$

where

$$\tilde{R} \triangleq \begin{bmatrix} R_1 & 0 \\ 0 & C_c^{\mathrm{T}} R_2 C_c \end{bmatrix}. \qquad (2.9)$$

Next, by introducing the performance variables

$$z(t) \triangleq E_1 x(t) + E_2 u(t) = \tilde{E} \tilde{x}(t), \qquad (2.10)$$

4

where $\tilde{E} \triangleq [E_1 \;\; E_2 C_c]$, and defining the transfer function from w to z by

$$\tilde{H}(s) \triangleq \tilde{E}(sI_{\tilde{n}} - \tilde{A})^{-1}\tilde{D},$$

where $\tilde{n} \triangleq n + n_c$, it can be shown that when $\tilde{A}$ is asymptotically stable, (2.8) is given by

$$J(A_c, B_c, C_c) = \|\tilde{H}(s)\|_2^2 \triangleq \frac{1}{2\pi} \int_{-\infty}^{\infty} \|\tilde{H}(j\omega)\|_F^2 d\omega.$$

For convenience we define the matrices $R_1 \triangleq E_1^T E_1$ and $R_2 \triangleq E_2^T E_2$ which are the H_2 weights for the state and control variables. Since $\tilde{A}$ is asymptotically stable, there exist nonnegative-definite matrices $\tilde{Q} \in \mathbf{R}^{\tilde{n} \times \tilde{n}}$ and $\tilde{P} \in \mathbf{R}^{\tilde{n} \times \tilde{n}}$ satisfying the closed-loop steady-state covariance equation and its dual, i.e.,

$$0 = \tilde{A}\tilde{Q} + \tilde{Q}\tilde{A}^T + \tilde{V}, \tag{2.11}$$

$$0 = \tilde{A}^T \tilde{P} + \tilde{P}\tilde{A} + \tilde{R}, \tag{2.12}$$

where

$$\tilde{V} \triangleq \begin{bmatrix} V_1 & 0 \\ 0 & B_c V_2 B_c^T \end{bmatrix}. \tag{2.13}$$

The cost functional (2.7) can now be expressed as

$$J(A_c, B_c, C_c) = \operatorname{tr} \tilde{Q}\tilde{R} = \operatorname{tr} \tilde{P}\tilde{V}. \tag{2.14}$$

Before presenting the main theorem we present a key lemma concerning nonnegative definite matrices and several definitions.

Lemma 2.1.[28]   Suppose $\hat{Q} \in \mathbf{R}^{n \times n}$ and $\hat{P} \in \mathbf{R}^{n \times n}$ are symmetric and nonnegative-definite and rank $\hat{Q}\hat{P} = n_c$. Then, the following statements hold:

(i) $\hat{Q}\hat{P}$ is diagonalizable and has nonnegative eigenvalues.

(ii) The $n \times n$ matrix

$$\tau \triangleq \hat{Q}\hat{P}(\hat{Q}\hat{P})^{\#}, \tag{2.15}$$

is idempotent, i.e., τ is an oblique projection and

$$\operatorname{rank} \tau = n_c. \tag{2.16}$$

Thus, if τ is given by (2.15), then there exists a nonsingular matrix $W \in \mathbf{R}^{n \times n}$ such that

$$\tau = W \begin{bmatrix} I_{n_c} & 0 \\ 0 & 0 \end{bmatrix} W^{-1}. \tag{2.17}$$

5

(*iii*) There exist $G, \Gamma \in \mathbf{R}^{n_c \times n}$ and nonsingular $M \in \mathbf{R}^{n_c \times n_c}$ such that

$$\hat{Q}\hat{P} = G^{\mathrm{T}} M \Gamma, \tag{2.18}$$

$$\Gamma G^{\mathrm{T}} = I_{n_c}. \tag{2.19}$$

(*iv*) If G, Γ, and M satisfy property (*iii*) then

$$\mathrm{rank}\ G = \mathrm{rank}\ \Gamma = \mathrm{rank}\ M = n_c, \tag{2.20}$$

$$(\hat{Q}\hat{P})^{\#} = G^{\mathrm{T}} M^{-1} \Gamma, \tag{2.21}$$

$$\tau = G^{\mathrm{T}} \Gamma, \tag{2.22}$$

$$\tau G^{\mathrm{T}} = G^{\mathrm{T}}, \Gamma \tau = \tau. \tag{2.23}$$

(*v*) The matrices G, Γ, and M satisfying property (*iii*) are unique except for a change of basis in $\mathbf{R}^{n_c}$, i.e., if G', Γ', and M' also satisfy property (*iv*), then there exists nonsingular $T_c \in \mathbf{R}^{n_c \times n_c}$ such that $G' = T_c^{\mathrm{T}} G, \Gamma' = T_c^{-1} \Gamma, M' = T_c^{-1} M T_c$. Furthermore, all such M are diagonalizable with positive eigenvalues.

(*vi*) Finally, if $\mathrm{rank}\ \hat{Q} = \mathrm{rank}\ \hat{P} = \mathrm{rank}\ \hat{Q}\hat{P} = n_c$, there exists a nonsingular transformation $W \in \mathbf{R}^{n \times n}$ such that

$$\hat{Q} = W \begin{bmatrix} \Omega & 0 \\ 0 & 0 \end{bmatrix} W^{\mathrm{T}}, \tag{2.24}$$

$$\hat{P} = W^{-\mathrm{T}} \begin{bmatrix} \Omega & 0 \\ 0 & 0 \end{bmatrix} W^{-1}, \tag{2.25}$$

where $\Omega \in \mathbf{R}^{n_c \times n_c}$ is diagonal and nonsingular. In addition,

$$\hat{Q} = \tau \hat{Q} = \hat{Q} \tau^{\mathrm{T}} = \tau \hat{Q} \tau^{\mathrm{T}}, \tag{2.26}$$

$$\hat{P} = \tau^{\mathrm{T}} \hat{P} = \hat{P} \tau = \tau^{\mathrm{T}} \hat{P} \tau. \tag{2.27}$$

Definition 2.1. A triple (G, M, Γ) satisfying property (*iii*) of Lemma 2.1 is a *projective factorization* of $\hat{Q}\hat{P}$.

Definition 2.2. A compensator (A_c, B_c, C_c) is an *extremal* of the optimal generalized fixed-order dynamic compensation problem if it satisfies the first order necessary conditions of optimality, i.e.,

$$\frac{\partial J}{\partial A_c} = 0, \quad \frac{\partial J}{\partial B_c} = 0, \quad \frac{\partial J}{\partial C_c} = 0,$$

6

where $J(A_c, B_c, C_c)$ is defined by (2.7).

Definition 2.3. A compensator (A_c, B_c, C_c) is an *admissible extremal* of the optimal generalized fixed-order dynamic compensation problem if it is an extremal and is also in $\mathcal{S}_c$, i.e., the closed-loop system is asymptotically stable.

Finally, for convenience in stating the main results we define

$$\bar{\Sigma} \triangleq C^{\mathrm{T}} V_2^{-1} C, \quad \Sigma \triangleq B R_2^{-1} B^{\mathrm{T}}. \tag{2.28}$$

Theorem 2.1.[20] Suppose (A_c, B_c, C_c) is an admissible extremal of the optimal fixed-order dynamic compensation problem. Then, there exist $n \times n$ nonnegative-definite matrices $P, Q, \hat{P}$, and $\hat{Q}$ such that $A_c, B_c,$ and C_c are given by

$$A_c = \Gamma(A - Q\bar{\Sigma} - \Sigma P + Q C^{\mathrm{T}} V_2^{-1} D R_2^{-1} B^{\mathrm{T}} P) G^{\mathrm{T}}, \tag{2.29}$$

$$B_c = \Gamma Q C^{\mathrm{T}} V_2^{-1}, \quad C_c = R_2^{-1} B^{\mathrm{T}} P G^{\mathrm{T}}, \tag{2.30}$$

for some projective factorization (G, M, Γ) of $\hat{Q}\hat{P}$ and such that the following conditions are satisfied:

$$0 = A^{\mathrm{T}} P + PA + R_1 - P\Sigma P + \tau_\perp^{\mathrm{T}} P\Sigma P \tau_\perp, \tag{2.31}$$

$$0 = AQ + QA^{\mathrm{T}} + V_1 - Q\bar{\Sigma}Q + \tau_\perp Q\bar{\Sigma}Q\tau_\perp^{\mathrm{T}}, \tag{2.32}$$

$$0 = (A - Q\bar{\Sigma})^{\mathrm{T}} \hat{P} + \hat{P}(A - Q\bar{\Sigma}) + P\Sigma P - \tau_\perp^{\mathrm{T}} P\Sigma P \tau_\perp, \tag{2.33}$$

$$0 = (A - \Sigma P)\hat{Q} + \hat{Q}(A - \Sigma P)^{\mathrm{T}} + Q\bar{\Sigma}Q - \tau_\perp Q\bar{\Sigma}Q\tau_\perp^{\mathrm{T}}, \tag{2.34}$$

$$\operatorname{rank} \hat{Q} = \operatorname{rank} \hat{P} = \operatorname{rank} \hat{Q}\hat{P} = n_c, \tag{2.35}$$

$$\tau = (\hat{Q}\hat{P})(\hat{Q}\hat{P})^{\#}, \quad \tau_\perp \triangleq I_n - \tau. \tag{2.36}$$

Furthermore, the minimal cost is given by

$$J(A_c, B_c, C_c) = \operatorname{tr}[PV_1 + Q(P\Sigma P - \tau_\perp^{\mathrm{T}} P\Sigma P \tau_\perp)], \tag{2.37}$$

or, equivalently,

$$J(A_c, B_c, C_c) = \operatorname{tr}[QR_1 + P(Q\bar{\Sigma}Q - \tau_\perp Q\bar{\Sigma}Q\tau_\perp^{\mathrm{T}})]. \tag{2.38}$$

7

Conversely, if there exist $n \times n$ nonnegative-definite matrices $P, Q, \hat{P}$, and $\hat{Q}$ satisfying (2.31)–(2.36) then the compensator (A_c, B_c, C_c) given by (2.29) and (2.30) is an extremal of the optimal fixed-order dynamic compensation problem. Furthermore, $\tilde{A}$ is asymptotically stable if and only if $(\tilde{A}, \tilde{E})$ is detectable (or, equivalently, $(\tilde{A}, \tilde{D})$ is stabilizable).

Remark 2.1. Partitioning $\tilde{Q}$ and $\tilde{P}$ given by (2.11) and (2.12), respectively, as

$$\tilde{Q} = \begin{bmatrix} Q_1 & Q_{12} \\ Q_{12}^{\mathrm{T}} & Q_2 \end{bmatrix}, Q_1 \in \mathbf{R}^{n \times n}, Q_2 \in \mathbf{R}^{n_c \times n_c}, \tag{2.39}$$

$$\tilde{P} = \begin{bmatrix} P_1 & P_{12} \\ P_{12}^{\mathrm{T}} & P_2 \end{bmatrix}, P_1 \in \mathbf{R}^{n \times n}, P_2 \in \mathbf{R}^{n_c \times n_c}, \tag{2.40}$$

it follows from Ref. 20 that $P, Q, \hat{P}$ and $\hat{Q}$ given by (2.31)–(2.36) can be expressed as

$$P = P_1 - P_{12} P_2^{-1} P_{12}^{\mathrm{T}}, \tag{2.41}$$

$$Q = Q_1 - Q_{12} Q_2^{-1} Q_{12}^{\mathrm{T}}, \tag{2.42}$$

$$\hat{P} = P_{12} P_2^{-1} P_{12}^{\mathrm{T}}, \tag{2.43}$$

and

$$\hat{Q} = Q_{12} Q_2^{-1} Q_{12}^{\mathrm{T}}, \tag{2.44}$$

respectively.

Theorem 2.1 shows that one can compute an optimal reduced-order controller by solving the set of coupled, modified Riccati and Lyapunov equations (2.31)–(2.34) subject to the rank condition constraints (2.35). One approach to find a solution of (2.31)–(2.34) is based on homotopy methods.

3. Homotopy Methods for the Solution of Nonlinear Algebraic Equations

A "homotopy" is a continuous deformation of one function into another. Over the past several years, homotopy or continuation methods (whose mathematical basis is algebraic topology and differential topology[29]) have received significant attention in the mathematics literature and have been applied successfully to several important problems[30–35]. Recently, the engineering literature has also begun to recognize the utility of these methods for engineering applications[36–45]. The purpose of this section is to provide a very brief description of homotopy methods for finding the solutions of nonlinear algebraic equations. The reader is referred to Ref. 35, 36 and 46 for additional details.

The basic problem is as follows. Given sets U and V contained in $\mathbf{R}^n$ and a mapping $F: U \to V$, find solutions $u \in U$ to satisfy

$$F(u) = 0. \tag{3.1}$$

Homotopy methods embed the problem (3.1) in a larger problem. In particular let $H: U \times [0,1] \to \mathbf{R}^n$ be such that:

1) $H(u, 1) = F(u)$.

2) There exists at least one known $u_0 \in \mathbf{R}^n$ which is a solution to $H(\cdot, 0) = 0$, i.e.,

$$H(u_0, 0) = 0. \tag{3.2}$$

3) There exists a continuous curve $(u(\lambda), \lambda)$ in $\mathbf{R}^n \times [0, 1]$ such that

$$H(u(\lambda), \lambda) = 0 \text{ for } \lambda \in [0, 1], \tag{3.3}$$

with

$$(u(0), 0) = (u_0, 0). \tag{3.4}$$

4) The curve $(u(\lambda), \lambda)$ is differentiable.

A homotopy algorithm then constructs a procedure to compute the actual curve such that the initial solution $u(0)$ is transformed to a desired solution $u(1)$ satisfying

$$0 = H(u(1), 1) = F(u(1)). \tag{3.5}$$

Now, differentiating $H(u(\lambda), \lambda) = 0$ with respect to λ yields Davidenko's differential equation

$$\frac{\partial H}{\partial u}\frac{du}{d\lambda} + \frac{\partial H}{\partial \lambda} = 0, \tag{3.6}$$

which together with $u(0) = u_0$, defines an initial value problem. The desired solution $u(1)$ is then obtained by numerical integration from 0 to 1. Some numerical integration schemes are described in Ref. 35 and 46.

9

4. A Homotopy Algorithm for H_2 Optimal Reduced-Order Control

This section begins by introducing a homotopy map based on the optimal projection equations. The construction of the initial point is then discussed in detail. Finally, the actual homotopy algorithm is presented.

4.1 The Homotopy Map

To define the homotopy map we assume that the plant matrices (A, B, C, D), the cost weighting matrices (R_1, R_2), the disturbance matrices (V_1, V_2) are functions of the homotopy parameter $\lambda \in [0, 1]$. In particular, the following is assumed:

$$\begin{bmatrix} A(\lambda) & B(\lambda) \\ C(\lambda) & D(\lambda) \end{bmatrix} = \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} + \lambda \left(\begin{bmatrix} A_f & B_f \\ C_f & D_f \end{bmatrix} - \begin{bmatrix} A_0 & B_0 \\ C_0 & D_0 \end{bmatrix} \right), \tag{4.1}$$

$$R_1(\lambda) = R_{1,0} + \lambda(R_{1,f} - R_{1,0}), \quad R_2(\lambda) = R_{2,0} + \lambda(R_{2,f} - R_{2,0}), \tag{4.2}$$

$$V_1(\lambda) = V_{1,0} + \lambda(V_{1,f} - V_{1,0}), \quad V_2(\lambda) = V_{2,0} + \lambda(V_{2,f} - V_{2,0}). \tag{4.3}$$

Note that the above equations imply that $A(0) = A_0$, $B(0) = B_0$, ..., $V_2(0) = V_{2,0}$, and that $A(1) = A_f$, $B(1) = B_f$, ..., $V_2(1) = V_{2,f}$. For notational simplification, we also define

$$\Sigma(\lambda) \triangleq B(\lambda)R_2^{-1}(\lambda)B^{\mathrm{T}}(\lambda), \quad \bar{\Sigma}(\lambda) \triangleq C^{\mathrm{T}}(\lambda)V_2^{-1}(\lambda)C(\lambda). \tag{4.4}$$

The homotopy formulation $0 = H((P, Q, \hat{P}, \hat{Q}), \lambda)$ is thus given by

$$\begin{aligned} 0 = {}& A(\lambda)^{\mathrm{T}}P(\lambda) + P(\lambda)A(\lambda) + R_1(\lambda) + \tau^{\mathrm{T}}(\lambda)P(\lambda)\Sigma(\lambda)P(\lambda)\tau(\lambda) \\ & - \tau^{\mathrm{T}}(\lambda)P(\lambda)\Sigma(\lambda)P(\lambda) - P(\lambda)\Sigma(\lambda)P(\lambda)\tau(\lambda), \end{aligned} \tag{4.5}$$

$$\begin{aligned} 0 = {}& A(\lambda)Q(\lambda) + Q(\lambda)A(\lambda)^{\mathrm{T}} + V_1(\lambda) + \tau(\lambda)Q(\lambda)\bar{\Sigma}(\lambda)Q(\lambda)\tau^{\mathrm{T}}(\lambda) \\ & - \tau(\lambda)Q(\lambda)\bar{\Sigma}(\lambda)Q(\lambda) - Q(\lambda)\bar{\Sigma}(\lambda)Q(\lambda)\tau^{\mathrm{T}}(\lambda), \end{aligned} \tag{4.6}$$

$$\begin{aligned} 0 = {}& (A(\lambda) - Q(\lambda)\bar{\Sigma}(\lambda))^{\mathrm{T}}\hat{P}(\lambda) + \hat{P}(\lambda)(A(\lambda) - Q(\lambda)\bar{\Sigma}(\lambda)) - \tau^{\mathrm{T}}(\lambda)P(\lambda)\Sigma(\lambda)P(\lambda)\tau(\lambda) \\ & + \tau^{\mathrm{T}}(\lambda)P(\lambda)\Sigma(\lambda)P(\lambda) + P(\lambda)\Sigma(\lambda)P(\lambda)\tau(\lambda), \end{aligned} \tag{4.7}$$

$$\begin{aligned} 0 = {}& (A(\lambda) - \Sigma(\lambda)P(\lambda))\hat{Q}(\lambda) + \hat{Q}(\lambda)(A(\lambda) - \Sigma(\lambda)P(\lambda))^{\mathrm{T}} - \tau(\lambda)Q(\lambda)\bar{\Sigma}(\lambda)Q(\lambda)\tau^{\mathrm{T}}(\lambda) \\ & + \tau(\lambda)Q(\lambda)\bar{\Sigma}(\lambda)Q(\lambda) + Q(\lambda)\bar{\Sigma}(\lambda)Q(\lambda)\tau^{\mathrm{T}}(\lambda), \end{aligned} \tag{4.8}$$

where

$$\operatorname{rank} \hat{Q}(\lambda) = \operatorname{rank} \hat{P}(\lambda) = \operatorname{rank} \hat{Q}(\lambda)\hat{P}(\lambda) = n_c, \tag{4.9}$$

10

$$\tau(\lambda) = \hat{Q}(\lambda)\hat{P}(\lambda)[\hat{Q}(\lambda)\hat{P}(\lambda)]^{\#}, \tag{4.10}$$

and $\lambda \in [0, 1]$.

4.2. Initial System Selection

Before describing the general logic and features of the homotopy algorithm for H_2 optimal reduced-order dynamic compensation, we first discuss the importance of the homotopy initialization and some guidelines for choosing the initial system matrices. It is assumed that the designer has supplied a set of system and weighting matrices, $S_f = (A_f, B_f, C_f, D_f, R_{1,f}, R_{2,f}, V_{1,f}, V_{2,f})$ describing the optimization problem whose solution is desired. In addition, it is assumed that the designer has chosen an initial set of related system matrices $S_0 = (A_0, B_0, C_0, D_0, R_{1,0}, R_{2,0}, V_{1,0}, V_{2,0})$ that has an easily obtained $(P_0, Q_0, \hat{P}_0, \hat{Q}_0)$ which is either a solution or a good approximation to the solution of the optimal projection equations corresponding to the initial system (i.e., (4.5)–(4.10) with $\lambda = 0$).

While in general homotopy methods ease the restriction that the starting point be close to some optimal of the optimization problem, the initial guess does affect the performance of the homotopy algorithm. For example, it is always possible to choose the initial system S_0 such that (A_0, B_0, C_0, D_0) is nonminimal with minimal dimension n_c. In this case, it is easy to show that the corresponding LQG compensator has minimal dimension $n_r \leq n_c$ and will usually have minimal dimension $n_r = n_c$. In the latter case, $(A_{c,0}, B_{c,0}, C_{c,0})$ is chosen as a minimal realization of the LQG compensator. However, we have seen experimentally that the corresponding homotopy can lead to failure of the homotopy algorithm. Similar observations have been made in Ref. 13. In particular, Ref. 13 shows that allowing the plant parameters to vary along the homotopy path can lead to the development of destabilizing controllers or path bifurcations.

The reason that the above type of homotopy would cause problems is somewhat intuitive since for a given λ, say $\lambda_1 \in [0, 1]$, a controller $(A_c(\lambda_1), B_c(\lambda_1), C_c(\lambda_1))$ that stabilizes the plant $(A(\lambda_1), B(\lambda_1), C(\lambda_1), D(\lambda_1))$ may not stabilize the plant $(A(\lambda_2), B(\lambda_2), C(\lambda_2), D(\lambda_2))$ for $\lambda_2 \neq \lambda_1$. Hence, below we present ways of constructing the initial system S_0 that does *not* require the plant parameters (A, B, C, D) to vary along the homotopy path. In this case, a controller that stabilizes the plant at λ_1 will also stabilize the plant at $\lambda_2 > \lambda_1$. This argument in itself does *not* ensure that at every step along the homotopy algorithm the controller design remains stabilizing. This is a subject that requires further research. It should be mentioned that another advantage of a homotopy that varies only the performance weights (R_1, R_2, V_1, V_2) is that the optimal controller

11

at each point is optimal with respect to the real nominal plant (A_f, B_f, C_f, D_f).

Now, we present two options for constructing S_0 as proposed in Ref. 15.

Option 1. One alternative is to choose A_0 to be stable (e.g., if A_f is stable, let $A_0 = A_f$ or if A_f is unstable, let $A_0 = A_f - \sigma I$ where σ is sufficiently large to ensure stability of A_0) and, as elaborated in Ref. 47 to choose either $(R_{1,0}, V_{2,0})$ or $(V_{1,0}, R_{2,0})$ where $R_{1,0} \geq 0$, $V_{1,0} \geq 0$, $R_{2,0} > 0$, and $V_{2,0} > 0$, as given below. (All other initial parameters are equal to their final values.)

(i) In a basis in which

$$A_0 = \begin{bmatrix} (A_0)_{11} & 0 \\ (A_0)_{21} & (A_0)_{22} \end{bmatrix}, \quad (A_0)_{11} \in \mathbf{R}^{n_c \times n_c}, \tag{4.11}$$

choose $R_{1,0}$ to be of the form

$$R_{1,0} = \begin{bmatrix} (R_{1,0})_{11} & 0 \\ 0 & 0 \end{bmatrix}, \quad (R_{1,0})_{11} \in \mathbf{R}^{n_c \times n_c}, \tag{4.12}$$

and for some positive scalar α choose

$$V_{2,0} = \alpha V_{2,f}. \tag{4.13}$$

(ii) In a basis in which

$$A_0 = \begin{bmatrix} (A_0)_{11} & (A_0)_{12} \\ 0 & (A_0)_{22} \end{bmatrix}, \quad (A_0)_{11} \in \mathbf{R}^{n_c \times n_c}, \tag{4.14}$$

choose $V_{1,0}$ to be of the form

$$V_{1,0} = \begin{bmatrix} (V_{1,0})_{11} & 0 \\ 0 & 0 \end{bmatrix}, \quad (V_{1,0})_{11} \in \mathbf{R}^{n_c \times n_c}, \tag{4.15}$$

and for some positive scalar α choose

$$R_{2,0} = \alpha R_{2,f}. \tag{4.16}$$

As discussed in Ref. 47, α appearing in (4.13) and (4.16) can always be chosen sufficiently large so that the corresponding LQG compensator is nearly nonminimal. In this case, $(A_{c,0}, B_{c,0}, C_{c,0})$ is easily obtained by reducing the LQG compensator to its (nearly) minimal realization using an appropriate technique such as balanced controller reduction[48]. Next, form the closed-loop system consisting of (A_0, B_0, C_0, D_0) and $(A_{c,0}, B_{c,0}, C_{c,0})$ and compute the initial guess $P_0, Q_0, \hat{P}_0$, and $\hat{Q}_0$ using (2.41)–(2.44), respectively. Since $(A_{c,0}, B_{c,0}, C_{c,0})$ is a close approximation to the minimal realization of the corresponding nearly nonminimal LQG compensator, $(P_0, Q_0, \hat{P}_0, \hat{Q}_0)$ is a good

12

approximation of the solution of the optimal projection equations corresponding to the initial system (i.e., (4.5)–(4.10) with $\lambda = 0$).

Option 2. A second alternative (which does not require A_0 to be stable) is based on the following experimental observation. The initial system can be chosen to correspond to a low authority control problem, e.g., one can choose

$$R_{2,0} = \alpha R_{2,f}, \quad V_{2,0} = \beta V_{2,f},$$

with α and β large and let all other initial system parameters equal their final values. In this case it has been observed that the reduced-order controller $(A_{c,r}, B_{c,r}, C_{c,r})$ obtained by suboptimal reduction of an LQG controller will often yield virtually the same cost as the LQG controller[49], hence indicating that $(A_{c,r}, B_{c,r}, C_{c,r})$ may be nearly optimal. In this case, we choose $(A_{c,0}, B_{c,0}, C_{c,0}) = (A_{c,r}, B_{c,r}, C_{c,r})$. (It should be noted that these observations are partially explained by the results in Ref. 47.) Then, follow the same procedure described in *option 1* to form the closed-loop system and compute the initial guess $(P_0, Q_0, \hat{P}_0, \hat{Q}_0)$.

4.3. The Derivative and Correction Equations

The homotopy presented next uses a predictor/corrector numerical integration scheme. The prediction step requires derivatives $(\dot{P}(\lambda), \dot{Q}(\lambda), \dot{\hat{P}}(\lambda), \dot{\hat{Q}}(\lambda))$, where $\dot{M} \triangleq \frac{dM}{d\lambda}$, while the correction step is based on using a Newton correction, denoted as $(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q})$. Before constructing the derivative and correction equations, we state the following useful properties about the derivatives of the contragredient transformation of $(\hat{Q}, \hat{P})$.

Using Lemma 2.1, equations (4.9) and (4.10) imply

$$\hat{Q}(\lambda) = W(\lambda)\Lambda(\lambda)W^{\mathrm{T}}(\lambda) = W_1(\lambda)\Omega(\lambda)W_1^{\mathrm{T}}(\lambda), \tag{4.17}$$

$$\hat{P}(\lambda) = U^{\mathrm{T}}(\lambda)\Lambda(\lambda)U(\lambda) = U_1(\lambda)\Omega(\lambda)U_1^{\mathrm{T}}(\lambda), \tag{4.18}$$

and

$$\tau(\lambda) = W(\lambda)\begin{bmatrix} I_{n_c} & 0 \\ 0 & 0 \end{bmatrix} U(\lambda) = W_1(\lambda)U_1^{\mathrm{T}}(\lambda), \tag{4.19}$$

where

$$W(\lambda) \triangleq [W_1(\lambda) \quad W_2(\lambda)], \quad W_1(\lambda) \in \mathbf{R}^{n \times n_c}, \quad W_2(\lambda) \in \mathbf{R}^{n \times (n-n_c)}, \tag{4.20}$$

$$U(\lambda) \triangleq \begin{bmatrix} U_1^{\mathrm{T}}(\lambda) \\ U_2^{\mathrm{T}}(\lambda) \end{bmatrix}, \quad U_1(\lambda) \in \mathbf{R}^{n \times n_c}, \quad U_2(\lambda) \in \mathbf{R}^{n \times (n-n_c)}, \tag{4.21}$$

$$U(\lambda) = W^{-1}(\lambda), \tag{4.22}$$

13

or, equivalently,

$$U(\lambda)W(\lambda) = I_n,$$ (4.23)

$$\Lambda(\lambda) \triangleq \begin{bmatrix} \Omega(\lambda) & 0 \\ 0 & 0 \end{bmatrix}, \quad \Omega(\lambda) \in \mathbf{R}^{n_c \times n_c},$$ (4.24)

and $\Omega(\lambda)$ is diagonal and positive-definite. For notational simplicity, we omit the argument λ in the subsequent equations.

The derivative equations, obtained by differentiating (4.5)–(4.8) with respect to λ, are given by

$$0 = A_P^T \dot{P} + \dot{P} A_P + R_P(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + R_P^T(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + V_P + V_P^T + \dot{R}_1,$$ (4.25)

$$0 = A_Q \dot{Q} + \dot{Q} A_Q^T + R_Q(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + R_Q^T(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + V_Q + V_Q^T + \dot{V}_1,$$ (4.26)

$$0 = A_w^T \dot{\hat{P}} + \dot{\hat{P}} A_w + R_{\hat{P}}(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + R_{\hat{P}}^T(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + V_{\hat{P}} + V_{\hat{P}}^T,$$ (4.27)

$$0 = A_u \dot{\hat{Q}} + \dot{\hat{Q}} A_u^T + R_{\hat{Q}}(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + R_{\hat{Q}}^T(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + V_{\hat{Q}} + V_{\hat{Q}}^T.$$ (4.28)

The correction equations, derived similarly by using the relationship between the Newton's method and a particular error homotopy, are given by

$$0 = A_P^T \Delta P + \Delta P A_P + R_P(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + R_P^T(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + E_P^*,$$ (4.29)

$$0 = A_Q \Delta Q + \Delta Q A_Q^T + R_Q(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + R_Q^T(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + E_Q^*,$$ (4.30)

$$0 = A_w^T \Delta \hat{P} + \Delta \hat{P} A_w + R_{\hat{P}}(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + R_{\hat{P}}^T(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + E_{\hat{P}}^*,$$ (4.31)

$$0 = A_u \Delta \hat{Q} + \Delta \hat{Q} A_u^T + R_{\hat{Q}}(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + R_{\hat{Q}}^T(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + V_{\hat{Q}}^*,$$ (4.32)

The detail derivation of (4.25)–(4.32) and the definitions of all the coefficients are described in the Appendix A. Comparing (4.29)–(4.32) with (4.25)–(4.28) reveals that the derivative and correction equations are identical in form. Thus, only one solution procedure would be required to solve both sets of equations. Each set of equations consist of four coupled Lyapunov equations. Since these equations are linear, using Kronecker algebra[51] they can be converted to the vector form $\mathcal{A}\chi = b$ where for (4.29)–(4.32) χ is a vector containing the independent elements of $\Delta P, \Delta Q, \Delta W_1, \Delta U_1$, and $\Delta \Omega$. $\mathcal{A}$ is then a square matrix of dimension $n(n+1) + (2nn_c + n_c^2)$. Inversion of $\mathcal{A}$ is hence very computationally intensive for even relatively small problems (e.g., $n = 20, n_c = 10$).

Fortunately, the coupling terms $R_{\Delta P}, R_{\Delta Q}, R_{\Delta \hat{P}}$, and $R_{\Delta \hat{Q}}$ which are linear functions of $(\Delta P, \Delta Q, \Delta \hat{P}. \Delta \hat{Q})$ or, equivalently, $(\Delta P, \Delta Q, \Delta W_1, \Delta U_1, \Delta \Omega)$ in (4.29)–(4.32), have relatively low

14

ranks. Hence, the technique of Ref. 21, which exploits this low rank property, can be used to *efficiently* solve equations (4.29)–(4.32) (or, equivalently, (4.25)–(4.28)). In particular, this solution procedure requires inversion of a square matrix of dimension $(2n + n_c)(m + l) + 1$ and to solve four sets of $(2n + n_c)(m + l)$ standard $n \times n$ Lyapunov equations, which has much less computational burden than the approach using Kronecker algebra as described in the previous paragraph. In comparison, the dimension of the homotopy Jacobian inverted in the minimal parameterization approach is $n_c(m + l)$ which is smaller than the characteristic dimension associated with this approach. However, the algorithms based on these minimal parameterization basis sometimes exhibit numerical ill-conditioning or can even fail due to the basis constraint. The details of the solution procedure are described in Appendix B.

Also, note that if the homotopy path exists, the solution to the coupled Lyapunov equations will be well-posed. Hence, the matrices $A_P, A_Q, A_u,$ and A_w in (4.25)–(4.32) will have the property that any two eigenvalues of a given matrix will not sum to zero.

4.4 Overview of the Homotopy Algorithm

Below, we present an outline of the homotopy algorithm. This algorithm describes a predictor/corrector numerical integration scheme. In order to force the rank conditions (4.9) of $\hat{Q}$ and $\hat{P}$ during intermediate steps, we use the following scheme to update $(P, Q, \hat{P}, \hat{Q})$ along the homotopy path. First, using (A.29)–(A.31) and (A.57)–(A.59) and the algorithms described in Appendix C, the prediction $(\dot{\hat{Q}}, \dot{\hat{P}})$ and correction $(\Delta\hat{Q}, \Delta\hat{P})$ are first converted to $(\dot{W}_1, \dot{U}_1, \dot{\Omega})$ and $(\Delta W_1, \Delta U_1, \Delta\Omega)$, respectively. Note that $\dot{\Omega}$ and $\Delta\Omega$ are forced to be $n_c \times n_c$ diagonal matrices with this formulation. Next, we update (P, Q, W_1, U_1, Ω) with these predictions/corrections. Finally, new $(\hat{Q}, \hat{P})$ are constructed with updated (W_1, U_1, Ω) using (4.17) and (4.18) and the rank conditions (4.9) are maintained.

There are several options to be chosen initially. These options are enumerated before presenting the actual algorithm. Note that each option corresponds to a particular flag being assigned some integer value.

4.4.1 Prediction Scheme Options

Here we use the notation $\lambda_0, \lambda_{-1},$ and λ_1 representing the values of λ at respectively the current point on the homotopy curve, the previous point, and the next point. Also, $\dot{M} \triangleq dM/d\lambda$ and $\theta(\lambda)$ is a vector representation of $(P(\lambda), Q(\lambda), W_1(\lambda), U_1(\lambda), \Omega(\lambda))$.

15

pred = 0. No prediction. This option assumes that $\theta(\lambda_1) = \theta(\lambda_0)$.

pred = 1. Linear prediction. This option assumes that $\theta(\lambda_1)$ is predicted using $\dot{\theta}(\lambda_0)$ and $\theta(\lambda_0)$. In particular,

$$\theta(\lambda_1) = \theta(\lambda_0) + (\lambda_1 - \lambda_0)\dot{\theta}(\lambda_0), \tag{4.33}$$

pred = 2. Cubic spline prediction. This prediction of $\theta(\lambda_1)$ requires $\theta(\lambda_0), \dot{\theta}(\lambda_0), \theta(\lambda_{-1})$, and $\dot{\theta}(\lambda_{-1})$. In particular,

$$\text{vec}[\theta(\lambda_1)] = a_0 + a_1\lambda_1 + a_2\lambda_1^2 + a_3\lambda_1^3,$$

where a_0, a_1, a_2, and a_3 are computed by solving

$$[a_0 \ a_1 \ a_2 \ a_3] \begin{bmatrix} 1 & 0 & 1 & 0 \\ \lambda_{-1} & 1 & \lambda_0 & 1 \\ \lambda_{-1}^2 & 2\lambda_{-1} & \lambda_0^2 & 2\lambda_0 \\ \lambda_{-1}^3 & 3\lambda_{-1}^2 & \lambda_0^3 & 3\lambda_0^2 \end{bmatrix} = \begin{bmatrix} \text{vec}[\theta(\lambda_{-1})] \\ \text{vec}[\dot{\theta}(\lambda_{-1})] \\ \text{vec}[\theta(\lambda_0)] \\ \text{vec}[\dot{\theta}(\lambda_0)] \end{bmatrix}^{\mathrm{T}}.$$

Note that if $\theta(\lambda_{-1})$ and $\dot{\theta}(\lambda_{-1})$ are not available (as occurs at the initial iteration of the homotopy algorithm), then $\theta(\lambda_1)$ is predicted using the linear prediction given by (4.33).

4.4.2. Basis Options for Solving the Coupled Lyapunov Equations

The main computational burden of the algorithm given below is the solution of the four coupled modified Lyapunov equations (4.25)–(4.28) or (4.29)–(4.32) at each prediction step or correction iteration. Efficient solutions of these equations, as described in Appendix B, makes the algorithm feasible for large scale systems. The most desired solution procedure is based on diagonalizing the coefficient matrices A_P, A_Q, A_w, and A_u of the coupled Lyapunov equations. This is usually possible. However, it is also possible that this diagonalization will be intractable for some points along the homotopy path. A numerical conditioning test is embedded in the program to determine whether the coefficient matrices are truly diagonalizable. If they are not, then the coupled Lyapunov equations are solved using the Schur decomposition. A second option relies exclusively on the Schur decomposition.

basis = 1. A_P, A_Q, A_w and A_u are diagonalized when solving (4.25)–(4.28) or (4.29)–(4.32).

basis = 2. A_P, A_Q, A_w and A_u are in Schur form when solving (4.25)–(4.28) or (4.29)–(4.32).

4.4.3. Outline of the Homotopy Algorithm

Step 1. Initialize *loop* = 0, $\lambda = 0$, $\Delta\lambda \in (0, 1]$, $S = S_0$, $(P, Q, \hat{P}, \hat{Q}) = (P_0, Q_0, \hat{P}_0, \hat{Q}_0)$.

16

Step 2. Let $loop = loop + 1$. If $loop = 1$, then go to Step 4.

Step 3. Advance the homotopy parameter λ and predict the corresponding $P(\lambda), Q(\lambda), \hat{P}(\lambda)$, and $\hat{Q}(\lambda)$ as follows.

    3a. Let $\lambda_0 = \lambda$.

    3b. Let $\lambda = \lambda_0 + \Delta\lambda$.

    3c. If $pred \geq 1$, then perform the next step to compute $\dot{P}(\lambda_0), \dot{Q}(\lambda_0), \dot{\hat{P}}(\lambda_0)$, and $\dot{\hat{Q}}(\lambda_0)$ according to (4.25)–(4.28). Else, let $P(\lambda) = P(\lambda_0), Q(\lambda) = Q(\lambda_0), \hat{P}(\lambda) = \hat{P}(\lambda_0)$, and $\hat{Q}(\lambda) = \hat{Q}(\lambda_0)$ and go to step (4), i.e., no prediction is performed.

    3d. Transform A_P, A_Q, A_w, and A_u into suitable matrix form according to the option defined by $basis$, then solve (4.25)–(4.28) as described in Appendix B.

    3e. Compute $(\dot{W}_1(\lambda_0), \dot{U}_1(\lambda_0), \dot{\Omega}(\lambda_0))$ from $(\dot{\hat{Q}}(\lambda_0), \dot{\hat{P}}(\lambda_0))$ by using (A.29)–(A.31) and the procedure described in Appendix C.

    3f. Predict $(P(\lambda), Q(\lambda), W_1(\lambda), U_1(\lambda), \Omega(\lambda))$ by using the option defined by $pred$.

    3g. Compute $(\hat{Q}(\lambda), \hat{P}(\lambda))$ from $(W_1(\lambda), U_1(\lambda), \Omega(\lambda))$ using (4.17) and (4.18).

Step 4. Correct the current approximations $P(\lambda^*), Q(\lambda^*), \hat{P}(\lambda^*)$, and $\hat{Q}(\lambda^*)$ as follows.

    4a. Compute the errors $(E_P^*, E_Q^*, E_{\hat{Q}}^*, E_{\hat{P}}^*)$ in the correction equations (A.38)–(A.41).

    4b. Transform A_P, A_Q, A_w, and A_u into suitable matrix form by using the option defined by $basis$, then solve (4.29)–(4.32) as described in the Appendix B for $\Delta P, \Delta Q, \Delta \hat{P}$, and $\Delta \hat{Q}$.

    4c. Compute $(\Delta W_1, \Delta U_1, \Delta \Omega)$, from $(\Delta \hat{Q}, \Delta \hat{P})$ by using (A.57)–(A.59) and the algorithms described in Appendix C.

    4d. Let

$$P(\lambda) \longleftarrow P(\lambda) + \Delta P, \quad Q(\lambda) \longleftarrow Q(\lambda) + \Delta Q,$$

$$W_1(\lambda) \longleftarrow W_1(\lambda) + \Delta W_1, \quad U_1(\lambda) \longleftarrow U_1(\lambda) + \Delta U_1, \quad \Omega(\lambda) \longleftarrow \Omega(\lambda) + \Delta \Omega.$$

    4e. Compute $(\hat{Q}(\lambda), \hat{P}(\lambda))$ from $(W_1(\lambda), U_1(\lambda), \Omega(\lambda))$ using (4.17) and (4.18).

    4f. Recompute the errors $(E_P^*, E_Q^*, E_{\hat{P}}^*, E_{\hat{Q}}^*)$ in the correction equations (A.38)–(A.41). If the

$$\max \left(\frac{\|E_P^*\|_A}{\|R_1(\lambda)\|_A}, \frac{\|E_Q^*\|_A}{\|V_1(\lambda)\|_A}, \frac{\|E_P^*\|_A}{\|\Sigma(\lambda)\|_A}, \frac{\|E_Q^*\|_A}{\|\Sigma(\lambda)\|_A} \right) < \delta^*, \text{ where } \delta^* \text{ is some preassigned correction}$$

tolerance, then set $\lambda_0 = \lambda$, and adjust next step size $\Delta\lambda$ according to the number of the correction steps required to converge before going to Step 3b. Else, if the number of corrections exceeds a preset limit, reduce $\Delta\lambda$ and go to Step 3b; otherwise, go to Step 4b.

Step 5. If $\lambda = 1$, then stop. Else, go to Step 2.

Note that the algorithm described above allows the step size $(\Delta\lambda)$ to vary dynamically depending on the speed of convergence which is gauged by the number of the correction steps. If the number is small (e.g., ≤ 3), we increase (e.g., double) the previous step size when computing the next step. If it takes many steps to converge (e.g., > 10), or does not converge, the step size is reduced (e.g., in half). δ^* in Step 4f is a preassigned correction error tolerance which can be assigned with two values in the program. One is the intermediate correction error tolerance which is used when $\lambda < 1$. The other value is the final correction error tolerance which is usually smaller and is used when $\lambda = 1$. The choice of the magnitudes of theses tolerances are problem dependent. In general, the intermediate correction tolerance is desired to be reasonably large to speed the homotopy curve following. However, the algorithm may fail to converge if these tolerances are too large. The final correction tolerance is usually small to ensure the accuracy of the final results.

5. Illustrative Numerical Examples

In this section we present two illustrative numerical examples that demonstrate the effectiveness of the proposed algorithm. For both examples, the MATLAB implementation of the homotopy algorithm to design the optimal reduced-order compensator was run on a 486, 33 MHz PC. The design parameters R_2 and V_2 were allowed to vary during the homotopy path.

First, we consider a control design for an axial beam with four disks attached as shown in Figure 5.1. This example was derived from a laboratory experiment[52] and has been considered in several subsequent publications[49,53-55]. The basic control objective for the four-disk problem is to control the angular displacement at the location of disk 1 using a torque input at the location of disk 3. It is also assumed that a torque disturbance enters the system at the location of disk 3.

The design philosophy adopted here is that the scaling q_2 of the nominal control weight $R_{2,0} = 1$ and the nominal sensor noise intensity $V_{2,0} = 1$ are simply "design knobs" used to determine the control authority. (Hence, $R_2(\lambda) = q_2(\lambda)R_{2,0}$ and $V_2(\lambda) = q_2(\lambda)V_{2,0}$.) Here, we consider the design of $2^{nd}, 4^{th}$, and 6th -order controllers for various authority levels.

Since at $q_2 = 10$, the $2^{nd}, 4^{th}$, and 6th reduced-order controllers by balancing are all good approximations of the corresponding optimal controller, respectively, we use this suboptimal controller to initialize the homotopy algorithm and deform this controller into the higher authority optimal controller corresponding to $q_2 = 1$. In each of the following passes, we increase the authority level by decreasing R_2 and V_2 by a factor of 10, i.e., $q_{2f} = 0.1q_{2,0}$, and at the end of each pass deform the initial optimal controller to the optimal controller corresponding to the higher authority level. This process is repeated for every reduced-order design. Figures 5.2–5.4 compare the optimal curves at various authority levels for an LQG controller, a reduced-order controller obtained by balancing and an optimal reduced- order controller. In each case, the optimal reduced-order controller performs better than the balanced controller as the authority is increased. Figure 5.5 compares the optimal controllers of various orders. This type of figure can be used in practice to determine the order of the controller to be implemented.

Control Authority $q_2(1)$	MegaFLOPs	RealTime (sec.)	Predictions &Corrections
10^{-1}	412	672	35
10^{-2}	407	727	35
10^{-3}	393	723	34
10^{-4}	274	478	24
10^{-5}	*	2990	120

Table 5.1. Run-Time statistics of Four Disk Example for $n_c = 4$

Table 5.1 shows some of the run time statistics for solving the 4th-order optimal compensator for this example at various control authorities. The cubic spline prediction option and the diagonal basis option were chosen for solving the coupled Lyapunov equations in this comparison. However, for $q_2(1) = 10^{-5}$ case, diagonalizing errors are significant and the basis option was switched to the Schur form.

The Frobenius norms of $P, Q, \hat{P}$, and $\hat{Q}$ are also recorded along the homotopy path and a typical results are shown in Figures 5.6–5.9 for the 4th-order controller design. It is interesting to note that as the control authority is increased beyond a certain level (e.g., for $n_c = 4, q_2 < 10^{-4}$) those values approach some stable limit as indicated in the figures. This is because $P, Q, \hat{P}$, and $\hat{Q}$ converge to fixed values as the control authority increases. It follows that the optimal reduced-order controller converges to a fixed value. This later phenomenon has been observed previously[15]. Furthermore, most of the prediction/correction steps indicated in Table 5.1 for this special case ($q_2(1) = 10^{-5}$) occur when λ is approaching 1 and $P, Q, \hat{P}$, and $\hat{Q}$ approach their final fixed values.

19

Algorithm efficiency as a function of the prediction options and basis options for solving the coupled Lyapunov equations has been studied in the context of a similar algorithm for H_2 optimal model reduction using the corresponding optimal projection equations[24]. It was seen that the algorithm is most efficient when using the cubic spline prediction and diagonalizing the coefficient matrices of the coupled Lyapunov equations. These conclusions also hold for the algorithm presented here.

The second example illustrates the design of an optimal reduced-order controller for a 17th order model of one of the single-input, single-output (SISO) transfer functions of the Active Control Technique Evaluation for Spacecraft (ACES) structure at NASA Marshall Space Flight Center [14,56]. The actuator and sensor are respectively a torque actuator and a collocated rate gyro. The model includes the actuator and sensor dynamics. A first order all-pass filter was appended to the model to approximate the computational delay associated with digital implementation.

Following the same approach, we design an 8th-order controller for this plant. Figure 5.10 shows the performance curves for authority levels corresponding to $q_2 \in (10^{-3}, 10^{-4}, \ldots, 10^{-7})$ and compare the optimal curves for an LQG controller, and an optimal reduced- order controller. For this special case, suboptimal reduced- order controllers obtained by balancing destabilize the closed-loop system when $q_2 \leq 10^{-5}$.

Table 5.2 shows some of the run time statistics for solving the 8th-order optimal compensator for this example at various control authorities. The cubic spline prediction option and the diagonal basis option were chosen for solving the coupled Lyapunov equations in this comparison. The "*" under the MegaFLOPs heading indicates that the MATLAB FLOP counter overflowed and so the FLOP data is unavailable.

Control Authority $q_2(1)$	MegaFLOPs	RealTime (sec.)	Predictions & Corrections
10^{-4}	*	4098	19
10^{-5}	*	9008	42
10^{-6}	*	4712	22
10^{-7}	*	2216	10

Table 5.2. Run-Time statistics of ACES Structure Example for $n_c = 8$

6. Conclusions

Gradient-based minimal parameterization homotopy algorithms for H_2 optimal reduced-order dynamic compensation[15] are computationally efficient in theory, but tend to experience numerical

ill-conditioning in practice due to the constraint on the controller basis. Hence, this paper has presented a new homotopy algorithm for the synthesis of H_2 optimal reduced-order compensators based on directly solving the optimal projection equations which characterize the optimal compensator. The resulting algorithm is usually more numerically robust than the gradient-based homotopy algorithms. However, the number of variables associated with this approach is $(2n + n_c) * (m + l)$ which is greater than the number of variables associated with minimal parameterization approach $(n_c(m+l))$. The two examples of the previous section demonstrate the effectiveness of the proposed algorithm.

Appendix A. Formulation of the Derivative and Correction Equations

Before deriving the derivative and correction equations (4.25)–(4.32), we state the following useful properties about the derivatives of the contragredient transformation of $(\hat{Q}, \hat{P})$.

Note that it follows from (4.17)-(4.19) that $\tau(\lambda)$ can be expressed as

$$\tau(\lambda) = \hat{Q}(\lambda)U^{\mathrm{T}}(\lambda)\Lambda^{\dagger}(\lambda)U(\lambda) = \hat{Q}(\lambda)U_1(\lambda)\Omega^{-1}(\lambda)U_1^{\mathrm{T}}(\lambda), \tag{A.1}$$

or

$$\tau(\lambda) = W(\lambda)\Lambda^{\dagger}(\lambda)W^{\mathrm{T}}(\lambda)\hat{P}(\lambda) = W_1(\lambda)\Omega^{-1}(\lambda)W_1^{\mathrm{T}}(\lambda)\hat{P}(\lambda), \tag{A.2}$$

where

$$\Lambda^{\dagger}(\lambda) = \begin{bmatrix} \Omega^{-1}(\lambda) & 0 \\ 0 & 0 \end{bmatrix}. \tag{A.3}$$

The representations of $\tau(\lambda)$ given by (A.1) and (A.2) are used below as a convenient way of expressing the derivative equations partially in terms of $\dot{\hat{Q}}(\lambda)$ and $\dot{\hat{P}}(\lambda)$ as opposed to expressing the derivative equations only in terms of $\dot{W}_1(\lambda), \dot{U}_1(\lambda)$, and $\dot{\Omega}(\lambda)$. For notational simplicity, we omit the argument λ in the subsequent equations.

Differentiating (A.1) or (A.2), gives the following expressions for $\dot{\tau}$

$$\dot{\tau} = \dot{\hat{Q}}U_1\Omega^{-1}U_1^{\mathrm{T}} + \hat{Q}(\dot{U}_1\Omega^{-1}U_1^{\mathrm{T}} + U_1\Omega^{-1}\dot{\Omega}\Omega^{-1}U_1^{\mathrm{T}} + U_1\Omega^{-1}\dot{U}_1^{\mathrm{T}}), \tag{A.4}$$

or

$$\dot{\tau} = W_1\Omega^{-1}W_1^{\mathrm{T}}\dot{\hat{P}} + (W_1\Omega^{-1}\dot{W}_1^{\mathrm{T}} + W_1\Omega^{-1}\dot{\Omega}\Omega^{-1}W_1^{\mathrm{T}} + \dot{W}_1\Omega^{-1}W_1^{\mathrm{T}})\hat{P}, \tag{A.5}$$

with

$$\frac{d\Omega^{-1}}{d\lambda} = -[\Omega^{-1}]^2\dot{\Omega} = -\Omega^{-1}\dot{\Omega}\Omega^{-1}, \tag{A.6}$$

21

since Ω is diagonal. Below, we derive the matrix equations that can be used to solve for the derivatives and corrections.

A.1. The Derivative Equations

Differentiating (4.5)–(4.8) with respect to λ and using (A.4)–(A.6), yields

$$0 = A_P^T \dot{P} + \dot{P} A_P + R_P(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + R_P^T(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + V_P + V_P^T + \dot{R}_1, \qquad (A.7)$$

$$0 = A_Q \dot{Q} + \dot{Q} A_Q^T + R_Q(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + R_Q^T(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + V_Q + V_Q^T + \dot{V}_1, \qquad (A.8)$$

$$0 = A_w^T \dot{\hat{P}} + \dot{\hat{P}} A_w + R_{\hat{P}}(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + R_{\hat{P}}^T(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + V_{\hat{P}} + V_{\hat{P}}^T, \qquad (A.9)$$

$$0 = A_u \dot{\hat{Q}} + \dot{\hat{Q}} A_u^T + R_{\hat{Q}}(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + R_{\hat{Q}}^T(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + V_{\hat{Q}} + V_{\hat{Q}}^T, \qquad (A.10)$$

where

$$A_P \triangleq A - \Sigma P \tau, \qquad (A.11)$$

$$A_w \triangleq A - Q\bar{\Sigma} + W_1 \Omega^{-1} W_1^T P \Sigma P (I_n - \tau), \qquad (A.12)$$

$$A_Q \triangleq A - \tau Q \bar{\Sigma}, \qquad (A.13)$$

$$A_u \triangleq A - \Sigma P + (I_n - \tau) Q \bar{\Sigma} Q U_1 \Omega^{-1} U_1^T, \qquad (A.14)$$

$$R_P \triangleq -\hat{P}\alpha(\dot{\hat{P}}, \dot{\hat{Q}}) P \Sigma P (I_n - \tau) - (I_n - \tau)^T P \Sigma \dot{P} \tau - \hat{P} W_1 \Omega^{-1} W_1^T P \Sigma P (I_n - \tau), \qquad (A.15)$$

$$R_Q \triangleq -\hat{Q}\gamma(\dot{\hat{P}}, \dot{\hat{Q}}) Q \bar{\Sigma} Q (I_n - \tau)^T - \tau \dot{Q} \bar{\Sigma} Q (I_n - \tau)^T - \hat{Q} U_1 \Omega^{-1} U_1^T Q \bar{\Sigma} Q (I_n - \tau)^T, \qquad (A.16)$$

$$R_{\hat{P}} \triangleq \hat{P}\alpha(\dot{\hat{P}}, \dot{\hat{Q}}) P \Sigma P (I_n - \tau) + (I_n - \tau)^T P \Sigma \dot{P} \tau + \tau^T P \Sigma \dot{P} - \bar{\Sigma} \dot{Q} \hat{P}, \qquad (A.17)$$

$$R_{\hat{Q}} \triangleq \hat{Q}\gamma(\dot{\hat{P}}, \dot{\hat{Q}}) Q \bar{\Sigma} Q (I_n - \tau)^T + \tau \dot{Q} \bar{\Sigma} Q (I_n - \tau)^T + \tau Q \bar{\Sigma} \dot{Q} - \Sigma \dot{P} \hat{Q}, \qquad (A.18)$$

$$\alpha(\dot{\hat{P}}, \dot{\hat{Q}}) \triangleq \dot{W}_1 \Omega^{-1} W_1^T - W_1 \Omega^{-1} \dot{\Omega} \Omega^{-1} W_1^T + W_1 \Omega^{-1} \dot{W}_1^T, \qquad (A.19)$$

$$\gamma(\dot{\hat{P}}, \dot{\hat{Q}}) \triangleq \dot{U}_1 \Omega^{-1} U_1^T - U_1 \Omega^{-1} \dot{\Omega} \Omega^{-1} U_1^T + U_1 \Omega^{-1} \dot{U}_1^T, \qquad (A.20)$$

$$V_P \triangleq \dot{A}^T P - (I_n - \frac{1}{2}\tau)]^T P \dot{\Sigma} P \tau, \qquad (A.21)$$

$$V_Q \triangleq \dot{A} Q - (I_n - \frac{1}{2}\tau) Q \dot{\bar{\Sigma}} Q \tau^T, \qquad (A.22)$$

$$V_{\hat{P}} \triangleq (\dot{A} + Q \dot{\bar{\Sigma}})^T \hat{P} + (I_n - \frac{1}{2}\tau)]^T P \dot{\Sigma} P \tau, \qquad (A.23)$$

$$V_{\hat{Q}} \triangleq (\dot{A} - \dot{\Sigma} P) \hat{Q} + (I_n - \frac{1}{2}\tau) Q \dot{\bar{\Sigma}} Q \tau^T. \qquad (A.24)$$

22

Note that it follows from (4.1)–(4.3),

$$\dot{A} = A_f - A_0, \quad \dot{B} = B_f - B_0, \quad \dot{C} = C_f - C_0, \qquad (A.25)$$

$$\dot{R}_1 = R_{1,f} - R_{1,0}, \quad \dot{R}_2 = R_{2,f} - R_{2,0}, \qquad (A.26)$$

$$\dot{V}_1 = V_{1,f} - V_{1,0}, \quad \dot{V}_2 = V_{2,f} - V_{2,0}, \qquad (A.27)$$

$$\dot{\Sigma} = \dot{B}R_2^{-1}B^{\mathrm{T}} - BR_2^{-2}\dot{R}_2 B^{\mathrm{T}} + BR_2^{-1}\dot{B}^{\mathrm{T}}, \quad \dot{\bar{\Sigma}} = \dot{C}^{\mathrm{T}}V_2^{-1}C + C^{\mathrm{T}}V_2^{-2}\dot{V}_2 C + C^{\mathrm{T}}V_2^{-1}\dot{C}. \qquad (A.28)$$

Next, differentiating (4.17) and (4.18), yields

$$\dot{Q} = \dot{W}_1 \Omega W_1^{\mathrm{T}} + W_1 \dot{\Omega} W_1^{\mathrm{T}} + W_1 \Omega \dot{W}_1^{\mathrm{T}}, \qquad (A.29)$$

$$\dot{P} = \dot{U}_1 \Omega U_1^{\mathrm{T}} + U_1 \dot{\Omega} U_1^{\mathrm{T}} + U_1 \Omega \dot{U}_1^{\mathrm{T}}. \qquad (A.30)$$

Furthermore, differentiating (4.23) with respect to λ gives

$$0 = \dot{U}W + U\dot{W} = \dot{U}_1^{\mathrm{T}}W_1 + U_1^{\mathrm{T}}\dot{W}_1. \qquad (A.31)$$

A.2. The Correction Equations

The correction equations are developed with λ at some fixed value, say λ^*. The derivation of the correction equations are based on the relationship between Newton's method and a particular homotopy. Below, we use the notation

$$\dot{f}(\theta) \triangleq \frac{\partial f}{\partial \theta}. \qquad (A.32)$$

Let $f : \mathbf{R}^n \to \mathbf{R}^n$ be C^1 continuous and consider the equation

$$0 = f(\theta). \qquad (A.33)$$

If $\theta^{(i)}$ is the current approximation to the solution of (A.33), then the Newton correction[50] $\Delta\theta$ is given by

$$\theta^{(i+1)} - \theta^{(i)} \triangleq \Delta\theta = -\dot{f}(\theta^{(i)})^{-1}e, \qquad (A.34)$$

where

$$e \triangleq f(\theta^{(i)}). \qquad (A.35)$$

Now, let $\theta^{(i)}$ be an approximation to θ satisfying (A.33). Then with e given by (A.35) construct the following homotopy to solve (A.33)

$$(1 - \beta)e = f(\theta(\beta)), \quad \beta \in [0, 1]. \qquad (A.36)$$

23

Note that at $\beta = 0$, (A.36) has solution $\theta(0) = \theta^{(i)}$ while $\theta(1)$ satisfies (A.33). Then differentiating (A.36) with respect to β gives

$$\frac{\partial \theta}{\partial \beta}\big|_{\beta=0} = -\dot{f}(\theta^{(i)})^{-1}e. \tag{A.37}$$

Remark A.1. Note that the Newton correction $\Delta\theta$ in (A.34) and the derivative $\frac{\partial\theta}{\partial\beta}\big|_{\beta=0}$ in (A.37) are identical. Hence, the Newton correction $\Delta\theta$ can be found by constructing a homotopy of the form (A.36) and solving for the resulting derivative $\frac{\partial\theta}{\partial\beta}\big|_{\beta=0}$. As seen below, this insight is particularly useful when deriving Newton corrections for equations that have a matrix structure.

Now, we use the insights of Remark A.1 to derive the equations that need to be solved for the Newton corrections $(\Delta P, \Delta Q, \Delta\hat{P}, \Delta\hat{Q})$, or, equivalently, $(\Delta P, \Delta Q, \Delta W_1, \Delta U_1, \Delta\Omega)$. We begin by recalling that λ is assumed to have some fixed value, say λ^*. Also, it is assumed that $(P^*, Q^*, \hat{P}^*, \hat{Q}^*)$ is the current approximation of $(P(\lambda^*), Q(\lambda^*), \hat{P}(\lambda^*), \hat{Q}(\lambda^*))$ and that $(E_P^*, E_Q^*, E_{\hat{P}}^*, E_{\hat{Q}}^*)$ are respectively the errors in equations (4.5)–(4.8) with $\lambda = \lambda^*$ and $(P(\lambda), Q(\lambda), \hat{P}(\lambda), \hat{Q}(\lambda))$ replaced by $(P^*, Q^*, \hat{P}^*, \hat{Q}^*)$, respectively.

Next, we form the homotopy

$$
\begin{aligned}
(1-\beta)E_P^* = {} & A^{\mathrm{T}}P(\beta) + P(\beta)A + R_1 + \tau^{\mathrm{T}}(\beta)P(\beta)\Sigma P(\beta)\tau(\beta) \\
& - \tau^{\mathrm{T}}(\beta)P(\beta)\Sigma P(\beta) - P(\beta)\Sigma P(\beta)\tau(\beta),
\end{aligned}
\tag{A.38}
$$

$$
\begin{aligned}
(1-\beta)E_Q^* = {} & AQ(\beta) + Q(\beta)A^{\mathrm{T}} + V_1 + \tau(\beta)Q(\beta)\bar{\Sigma}Q(\beta)\tau^{\mathrm{T}}(\beta) \\
& - \tau(\beta)Q(\beta)\bar{\Sigma}Q(\beta) - Q(\beta)\bar{\Sigma}Q(\beta)\tau^{\mathrm{T}}(\beta),
\end{aligned}
\tag{A.39}
$$

$$
\begin{aligned}
(1-\beta)E_{\hat{P}}^* = {} & (A - Q(\beta)\bar{\Sigma})^{\mathrm{T}}\hat{P}(\beta) + \hat{P}(\beta)(A - Q(\beta)\bar{\Sigma}) - \tau^{\mathrm{T}}(\beta)P(\beta)\Sigma P(\beta)\tau(\beta) \\
& + \tau^{\mathrm{T}}(\beta)P(\beta)\Sigma P(\beta) + P(\beta)\Sigma P(\beta)\tau(\beta),
\end{aligned}
\tag{A.40}
$$

$$
\begin{aligned}
(1-\beta)E_{\hat{Q}}^* = {} & (A(\beta) - \Sigma P(\beta))\hat{Q}(\beta) + \hat{Q}(\beta)(A - \Sigma P(\beta))^{\mathrm{T}} - \tau(\beta)Q(\beta)\bar{\Sigma}(\beta)Q(\beta)\tau^{\mathrm{T}}(\beta) \\
& + \tau(\beta)Q(\beta)\bar{\Sigma}(\beta)Q(\beta) + Q(\beta)\bar{\Sigma}(\beta)Q(\beta)\tau^{\mathrm{T}}(\beta).
\end{aligned}
\tag{A.41}
$$

Here, $(A, \Sigma, \bar{\Sigma}, R_1, V_1) = (A(\lambda^*), \Sigma(\lambda^*), \bar{\Sigma}(\lambda^*), R_1(\lambda^*), V_1(\lambda^*))$ are assumed to be evaluated at $\lambda = \lambda^*$ and at $\beta = 0$, $(P(0), Q(0), \hat{P}(0), \hat{Q}(0), \tau(0))$ are the current approximations. Differentiating (A.38)–(A.41) with respect to β, noting the identity of (A.4)–(A.6) with $\dot{\tau}$ now representing $\frac{d\tau}{d\beta}$, and using Remark A.1 to make the replacements

$$\Delta P \triangleq \frac{dP}{d\beta}\big|_{\beta=0}, \quad \Delta Q \triangleq \frac{dQ}{d\beta}\big|_{\beta=0}, \quad \Delta\hat{P} \triangleq \frac{d\hat{P}}{d\beta}\big|_{\beta=0}, \quad \Delta\hat{Q} \triangleq \frac{d\hat{Q}}{d\beta}\big|_{\beta=0}, \tag{A.42}$$

yields

$$0 = A_P^{\mathrm{T}}\Delta P + \Delta P A_P + R_P(\Delta P, \Delta Q, \Delta\hat{P}, \Delta\hat{Q}) + R_P^{\mathrm{T}}(\Delta P, \Delta Q, \Delta\hat{P}, \Delta\hat{Q}) + E_P^*, \tag{A.43}$$

24

$$0 = A_Q \Delta Q + \Delta Q A_Q^{\mathrm{T}} + R_Q(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + R_Q^{\mathrm{T}}(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + E_Q^{\bullet}, \qquad (A.44)$$

$$0 = A_w^{\mathrm{T}} \Delta \hat{P} + \Delta \hat{P} A_w + R_{\hat{P}}(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + R_{\hat{P}}^{\mathrm{T}}(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + E_{\hat{P}}^{\bullet}, \qquad (A.45)$$

$$0 = A_u \Delta \hat{Q} + \Delta \hat{Q} A_u^{\mathrm{T}} + R_{\hat{Q}}(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + R_{\hat{Q}}^{\mathrm{T}}(\Delta P, \Delta Q, \Delta \hat{P}, \Delta \hat{Q}) + V_{\hat{Q}}^{\bullet}, \qquad (A.46)$$

where

$$A_P \triangleq A - \Sigma P \tau, \qquad (A.47)$$

$$A_w \triangleq A - Q\bar{\Sigma} + W_1 \Omega^{-1} W_1^{\mathrm{T}} P \Sigma P (I_n - \tau), \qquad (A.48)$$

$$A_Q \triangleq A - \tau Q \bar{\Sigma}, \qquad (A.49)$$

$$A_u \triangleq A - \Sigma P + (I_n - \tau) Q \bar{\Sigma} Q U_1 \Omega^{-1} U_1^{\mathrm{T}}, \qquad (A.50)$$

$$R_P \triangleq -\hat{P}\alpha(\Delta\hat{P}, \Delta\hat{Q}) P\Sigma P(I_n - \tau) - (I_n - \tau)^{\mathrm{T}} P\Sigma \Delta P\tau - \Delta\hat{P} W_1 \Omega^{-1} W_1^{\mathrm{T}} P\Sigma P(I_n - \tau), \quad (A.51)$$

$$R_Q \triangleq -\hat{Q}\gamma(\Delta\hat{P}, \Delta\hat{Q}) Q\bar{\Sigma}Q(I_n - \tau)^{\mathrm{T}} - \tau\Delta Q\bar{\Sigma}Q(I_n - \tau)^{\mathrm{T}} - \Delta\hat{Q} U_1 \Omega^{-1} U_1^{\mathrm{T}} Q\bar{\Sigma}Q(I_n - \tau)^{\mathrm{T}}, \quad (A.52)$$

$$R_{\hat{P}} \triangleq \hat{P}\alpha(\Delta\hat{P}, \Delta\hat{Q}) P\Sigma P(I_n - \tau) + (I_n - \tau)^{\mathrm{T}} P\Sigma\Delta P\tau + \tau^{\mathrm{T}} P\Sigma\Delta P - \bar{\Sigma}\Delta Q\hat{P}, \qquad (A.53)$$

$$R_{\hat{Q}} \triangleq \hat{Q}\gamma(\Delta\hat{P}, \Delta\hat{Q}) Q\bar{\Sigma}Q(I_n - \tau)^{\mathrm{T}} + \tau\Delta Q\bar{\Sigma}Q(I_n - \tau)^{\mathrm{T}} + \tau Q\bar{\Sigma}\Delta Q - \Sigma\Delta P\hat{Q}, \qquad (A.54)$$

$$\alpha(\Delta\hat{P}, \Delta\hat{Q}) \triangleq \Delta W_1 \Omega^{-1} W_1^{\mathrm{T}} - W_1 \Omega^{-1} \Delta\Omega \Omega^{-1} W_1^{\mathrm{T}} + W_1 \Omega^{-1} \Delta W_1^{\mathrm{T}} W_1^{\mathrm{T}}, \qquad (A.55)$$

$$\gamma(\Delta\hat{P}, \Delta\hat{Q}) \triangleq \Delta U_1 \Omega^{-1} U_1^{\mathrm{T}} - U_1 \Omega^{-1} \Delta\Omega \Omega^{-1} U_1^{\mathrm{T}} + U_1 \Omega^{-1} \Delta U_1^{\mathrm{T}}. \qquad (A.56)$$

Next, replacing λ with β in (4.17), (4.18), and (4.23) and differentiating them with respect to β, the following equations are derived,

$$\Delta \hat{Q} = \Delta W_1 \Omega W_1^{\mathrm{T}} + W_1 \Delta\Omega W_1^{\mathrm{T}} + W_1 \Omega \Delta W_1^{\mathrm{T}}, \qquad (A.57)$$

$$\Delta \hat{P} = \Delta U_1 \Omega U_1^{\mathrm{T}} + U_1 \Delta\Omega U_1^{\mathrm{T}} + U_1 \Omega \Delta U_1^{\mathrm{T}}, \qquad (A.58)$$

$$0 = \Delta U_1^{\mathrm{T}} W_1 + U_1^{\mathrm{T}} \Delta W_1, \qquad (A.59)$$

where

$$\Delta U_1 \triangleq \frac{dU}{d\beta}\Big|_{\beta=0}, \quad \Delta W_1 \triangleq \frac{dW}{d\beta}\Big|_{\beta=0}, \quad \Delta\Omega_1 \triangleq \frac{d\Omega}{d\beta}\Big|_{\beta=0}.$$

Appendix B. Efficient Computation of the Solution to the Prediction and Correction Equations.

This appendix presents a solution procedure using Ref. 21 for efficiently solving the prediction equations (4.25)–(4.28) and the correction equations (4.29)–(4.32). We commence by recognizing that (4.25)–(4.28) and (4.29)–(4.32) have the following generic form:

$$0 = A_P^{\mathrm{T}}\dot{P} + \dot{P}A_P + \mathcal{F}_a(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + F_P, \qquad (B.1)$$

$$0 = A_Q\dot{Q} + \dot{Q}A_Q^{\mathrm{T}} + \mathcal{F}_b(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + F_Q, \qquad (B.2)$$

$$0 = A_w^{\mathrm{T}}\dot{\hat{P}} + \dot{\hat{P}}A_w + \mathcal{F}_c(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + F_{\dot{P}}, \qquad (B.3)$$

$$0 = A_u\dot{\hat{Q}} + \dot{\hat{Q}}A_u^{\mathrm{T}} + \mathcal{F}_d(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + F_{\dot{Q}}, \qquad (B.4)$$

where the linear operators $\mathcal{F}_a : \mathrm{X}_1^4\mathrm{S}^{n\times n} \to \mathrm{S}^{n\times n}$, $\mathcal{F}_b : \mathrm{X}_1^4\mathrm{S}^{n\times n} \to \mathrm{S}^{n\times n}$, $\mathcal{F}_c : \mathrm{X}_1^4\mathrm{S}^{n\times n} \to \mathrm{S}^{n\times n}$, and $\mathcal{F}_d : \mathrm{X}_1^4\mathrm{S}^{n\times n} \to \mathrm{S}^{n\times n}$, are defined by

$$\mathcal{F}_a(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) \triangleq R_P(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + R_P^{\mathrm{T}}(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}), \qquad (B.5)$$

$$\mathcal{F}_b(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) \triangleq R_Q(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + R_Q^{\mathrm{T}}(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}), \qquad (B.6)$$

$$\mathcal{F}_c(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) \triangleq R_{\dot{P}}(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + R_{\dot{P}}^{\mathrm{T}}(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}), \qquad (B.7)$$

$$\mathcal{F}_d(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) \triangleq R_{\dot{Q}}(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) + R_{\dot{Q}}^{\mathrm{T}}(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}), \qquad (B.8)$$

and F_P, F_Q, F_{HP} and $F_{\dot{Q}}$ are constant forcing terms. It is easy to verify that $(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}, \mathcal{F}_a, \mathcal{F}_b, \mathcal{F}_c, \mathcal{F}_d, F_P, F_Q, F_{\dot{P}}, F_{\dot{Q}})$ in the above equations represents $(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}, R_P + R_P^{\mathrm{T}}, R_Q + R_Q^{\mathrm{T}}, R_{\dot{P}} + R_{\dot{P}}^{\mathrm{T}}, R_{\dot{Q}} + R_{\dot{Q}}^{\mathrm{T}}, V_P + V_P^{\mathrm{T}}, V_Q + V_Q^{\mathrm{T}}, V_{\dot{P}} + V_{\dot{P}}^{\mathrm{T}}, V_{\dot{Q}} + V_{\dot{Q}}^{\mathrm{T}})$ in (4.25)–(4.28) and $(\Delta P, \Delta Q, \Delta\hat{P}, \Delta\hat{Q}, R_{\Delta P} + R_{\Delta P}^{\mathrm{T}}, R_{\Delta Q} + R_{\Delta Q}^{\mathrm{T}}, R_{\Delta\hat{P}} + R_{\Delta\hat{P}}^{\mathrm{T}}, R_{\Delta\hat{Q}} + R_{\Delta\hat{Q}}^{\mathrm{T}}, E_P^{\star}, E_Q^{\star}, E_{\dot{P}}^{\star}, E_{\dot{Q}}^{\star})$ in (4.29)–(4.32), respectively. Our goal now is to find for some integer $m_{\star}$ (as small as possible) linear operators $\phi : \mathrm{X}_1^4\mathrm{S}^{n\times n} \to \mathrm{R}^{m_{\star}}$, $\mathcal{G}_a : \mathrm{R}^{m_{\star}} \to \mathrm{S}^{n\times n}$, $\mathcal{G}_b : \mathrm{R}^{m_{\star}} \to \mathrm{S}^{n\times n}$, $\mathcal{G}_c : \mathrm{R}^{m_{\star}} \to \mathrm{S}^{n\times n}$, $\mathcal{G}_d : \mathrm{R}^{m_{\star}} \to \mathrm{S}^{n\times n}$, such that

$$\mathcal{F}_a(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) \triangleq \mathcal{G}_a(\phi(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}})), \qquad (B.9)$$

$$\mathcal{F}_b(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) \triangleq \mathcal{G}_b(\phi(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}})), \qquad (B.10)$$

$$\mathcal{F}_c(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) \triangleq \mathcal{G}_c(\phi(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}})), \qquad (B.11)$$

$$\mathcal{F}_d(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}}) \triangleq \mathcal{G}_d(\phi(\dot{P}, \dot{Q}, \dot{\hat{P}}, \dot{\hat{Q}})). \qquad (B.12)$$

Next, define the following linear operators

$$\phi_1(\dot{\hat{P}}, \dot{\hat{Q}}) \triangleq \mathrm{vec}\ U_1^{\mathrm{T}}\alpha(\dot{\hat{P}}, \dot{\hat{Q}})PB, \qquad (B.13)$$

26

$$\phi_2(\dot{P}) \triangleq \text{vec } B^T \dot{P}, \tag{B.14}$$

$$\phi_3(\dot{P}) \triangleq \text{vec } \dot{P} W_1 \Omega^{-1} W_1^T PB, \tag{B.15}$$

$$\phi_4(\dot{P}, \dot{Q}) \triangleq \text{vec } W_1^T \gamma(\dot{P}, \dot{Q}) QC^T, \tag{B.16}$$

$$\phi_5(\dot{Q}) \triangleq \text{vec } C\dot{Q}, \tag{B.17}$$

$$\phi_6(\dot{Q}) \triangleq \text{vec } \dot{Q} U_1 \Omega^{-1} U_1^T QC^T, \tag{B.18}$$

and

$$\phi \triangleq [\phi_1^T \ \phi_2^T \ \phi_3^T \ \phi_4^T \ \phi_5^T \ \phi_6^T \ ]^T, \tag{B.19}$$

where

$$\alpha(\dot{P}, \dot{Q}) = \dot{W}_1 \Omega^{-1} W_1^T - W_1 \Omega^{-1} \dot{\Omega} \Omega^{-1} W_1^T + W_1 \Omega^{-1} \dot{W}_1^T, \tag{B.20}$$

$$\gamma(\dot{P}, \dot{Q}) = \dot{U}_1 \Omega^{-1} U_1^T - U_1 \Omega^{-1} \dot{\Omega} \Omega^{-1} U_1^T + U_1 \Omega^{-1} \dot{U}_1^T, \tag{B.21}$$

are given by (A.19) and (A.20), respectively. It follows from (B.13)–(B.18) that $\phi_1(\cdot) \in \mathbf{R}^{m_1}$, $\phi_2(\cdot) \in \mathbf{R}^{m_2}$, $\phi_3(\cdot) \in \mathbf{R}^{m_3}$, $\phi_4(\cdot) \in \mathbf{R}^{m_4}$, $\phi_5(\cdot) \in \mathbf{R}^{m_5}$, $\phi_6(\cdot) \in \mathbf{R}^{m_6}$, and $\phi(\cdot) \in \mathbf{R}^{m_*}$ where $m_1 = n_c m$, $m_2 = nm$, $m_3 = nm$, $m_4 = n_c l$, $m_5 = nl$, $m_6 = nl$, and

$$m_* \triangleq \sum_{i=1}^{6} m_i = (2n + n_c)(m + l). \tag{B.22}$$

Note that it is assumed in (B.20) and (B.21) that $\dot{W}_1, \dot{U}_1$, and $\dot{\Omega}$ are obtained from (A.29)–(A.31). A procedure to compute $\dot{W}_1, \dot{U}_1$, and $\dot{\Omega}$ given $\dot{Q}$ and $\dot{P}$ is presented in Appendix C.

Now, note that (4.19)–(4.22) imply that

$$U_1^T W_1 = I_{n_c}, \quad U_1^T W_2 = 0, \quad U_2^T W_1 = 0, \quad I_n - \tau = W_2 U_2^T. \tag{B.23}$$

Thus, using (4.17)–(4.19) and (B.25), we can rewrite (A.15)–(A.18) as

$$R_P = F_{a1} \text{ vec}^{-1}(\phi_1) G_{a1} + F_{a2} \text{ vec}^{-1}(\phi_2) G_{a2} + F_{a3} \text{ vec}^{-1}(\phi_3) G_{a3}, \tag{B.24}$$

$$R_Q = F_{b1} \text{ vec}^{-1}(\phi_4) G_{b1} + F_{b2} \text{ vec}^{-1}(\phi_5) G_{b2} + F_{b3} \text{ vec}^{-1}(\phi_6) G_{b3}, \tag{B.25}$$

$$R_{\dot{P}} = F_{c1} \text{ vec}^{-1}(\phi_1) G_{c1} + F_{c2} \text{ vec}^{-1}(\phi_5) G_{c2} + F_{c3} \text{ vec}^{-1}(\phi_2) G_{c3} + F_{c4} \text{ vec}^{-1}(\phi_2) G_{c4}, \tag{B.26}$$

$$R_{\dot{Q}} = F_{d1} \text{ vec}^{-1}(\phi_4) G_{d1} + F_{d2} \text{ vec}^{-1}(\phi_5) G_{d2} + F_{d3} \text{ vec}^{-1}(\phi_5) G_{d3} + F_{d4} \text{ vec}^{-1}(\phi_2) G_{d4}, \tag{B.27}$$

where

$$F_{a1} = -U_1\Omega, \quad G_{a1} = R_2^{-1}B^{\mathrm{T}}PW_2U_2^{\mathrm{T}}, \quad F_{a2} = -U_2W_2^{\mathrm{T}}PBR_2^{-1}, \quad G_{a2} = W_1U_1^{\mathrm{T}},$$

$$F_{a3} = -I_n, \quad G_{a3} = R_2^{-1}B^{\mathrm{T}}PW_2U_2^{\mathrm{T}}, \tag{B.28}$$

$$F_{b1} = -W_1\Omega, \quad G_{b1} = V_2^{-1}CQU_2W_2^{\mathrm{T}}, \quad F_{b2} = -W_2U_2^{\mathrm{T}}QC^{\mathrm{T}}V_2^{-1}, \quad G_{b2} = U_1W_1^{\mathrm{T}},$$

$$F_{b3} = -I_n, \quad G_{b3} = V_2^{-1}CQU_2W_2^{\mathrm{T}}, \tag{B.29}$$

$$F_{c1} = U_1\Omega, \quad G_{c1} = R_2^{-1}B^{\mathrm{T}}PW_2U_2^{\mathrm{T}}, \quad F_{c2} = -C^{\mathrm{T}}V_2^{-1}, \quad G_{c2} = U_1\Omega U_1^{\mathrm{T}},$$

$$F_{c3} = U_2W_2^{\mathrm{T}}PBR_2^{-1}, \quad G_{c3} = W_1U_1^{\mathrm{T}}, \quad F_{c4} = U_1W_1^{\mathrm{T}}PBR_2^{-1}, \quad G_{c4} = I_n, \tag{B.30}$$

$$F_{d1} = W_1\Omega, \quad G_{d1} = V_2^{-1}CQU_2W_2^{\mathrm{T}}, \quad F_{d2} = W_2U_2^{\mathrm{T}}QC^{\mathrm{T}}V_2^{-1}, \quad G_{d2} = U_1W_1^{\mathrm{T}},$$

$$F_{d3} = W_1U_1^{\mathrm{T}}QC^{\mathrm{T}}V_2^{-1}, \quad G_{d3} = I_n, \quad F_{d4} = -BR_2^{-1}, \quad G_{d4} = W_1\Omega W_1^{\mathrm{T}}. \tag{B.31}$$

Next, define the following shaping matrices

$$E_{m_1} \triangleq [I_{m_1}\ 0], \quad E_{m_2} \triangleq [0_{m_2 \times m_1}\ I_{m_2}0], \quad E_{m_3} \triangleq [0_{m_3 \times (m_1+m_2)}\ I_{m_3}\ 0],$$

$$E_{m_4} \triangleq [0\ I_{m_4}\ 0_{m_4 \times (m_5+m_6)}], \quad E_{m_5} \triangleq [0\ I_{m_5}\ 0_{m_5 \times m_6}], \quad E_{m_6} \triangleq [0\ I_{m_6}], \tag{B.32}$$

where $E_{m_i} \in \mathbf{R}^{m_i \times m}$ for $i = (1, 2, ..., 6)$. Now, using these shaping matrices, we define

$$\mathcal{G}_{a1}(z) \triangleq F_{a1}\ \mathrm{vec}^{-1}(E_{m_1}z)G_{a1} + F_{a2}\ \mathrm{vec}^{-1}(E_{m_2}z)G_{a2} + F_{a3}\ \mathrm{vec}^{-1}(E_{m_3}z)G_{a3},$$

$$\mathcal{G}_{b1}(z) \triangleq F_{b1}\ \mathrm{vec}^{-1}(E_{m_4}z)G_{b1} + F_{b2}\ \mathrm{vec}^{-1}(E_{m_5}z)G_{b2} + F_{b3}\ \mathrm{vec}^{-1}(E_{m_6}z)G_{b3},$$

$$\mathcal{G}_{c1}(z) \triangleq F_{c1}\ \mathrm{vec}^{-1}(E_{m_1}z)G_{c1} + F_{c2}\ \mathrm{vec}^{-1}(E_{m_5}z)G_{c2} + F_{c3}\ \mathrm{vec}^{-1}(E_{m_2}z)G_{c3}$$
$$+ F_{c4}\ \mathrm{vec}^{-1}(E_{m_2}z)G_{c4},$$

$$\mathcal{G}_{d1}(z) \triangleq F_{d1}\ \mathrm{vec}^{-1}(E_{m_4}z)G_{d1} + F_{d2}\ \mathrm{vec}^{-1}(E_{m_5}z)G_{d2} + F_{d3}\ \mathrm{vec}^{-1}(E_{m_5}z)G_{d3}$$
$$+ F_{d4}\ \mathrm{vec}^{-1}(E_{m_2}z)G_{d4}, \tag{B.33}$$

and

$$\mathcal{G}_a \triangleq \mathcal{G}_{a1} + \mathcal{G}_{a1}^{\mathrm{T}}, \quad \mathcal{G}_b \triangleq \mathcal{G}_{b1} + \mathcal{G}_{b1}^{\mathrm{T}}, \quad \mathcal{G}_c \triangleq \mathcal{G}_{c1} + \mathcal{G}_{c1}^{\mathrm{T}}, \quad \mathcal{G}_d \triangleq \mathcal{G}_{d1} + \mathcal{G}_{d1}^{\mathrm{T}}. \tag{B.34}$$

It follows from (B.24)–(B.27) that $\mathcal{G}_a, \mathcal{G}_b, \mathcal{G}_c$, and $\mathcal{G}_d$ satisfy (B.9)–(B.12), respectively.

Now, with the above definitions, the solution procedure for a set of coupled, modified Lyapunov equations described in Ref. 21 is applied to solve for $(\dot{P}, \dot{Q}, \dot{P}, \dot{Q})$ in (B.1)–(B.4). With the above formulation, the efficiency of the coupled Lyapunov equations solver is realized by exploiting the

low rank properties of the coupling terms. In particular, to solve the prediction equations (B.1)–(B.4), the primary computational burden is to invert a matrix of $m_* \times m_*$, or, equivalently, $(2n + n_c)(m + l) \times (2n + n_c)(m + l)$ and to solve four sets of $m_* + 1$ standard $n \times n$ Lyapunov equations, with each set having coefficient matrix of A_P, A_Q, A_w, and A_u, respectively.

In comparison, by using Kronecker algebra[50], (B.1)–(B.4) can be converted to the vector form $\mathcal{A}\chi = b$ where χ is a vector consisting of the independent elements of $\dot{P}, \dot{Q}, \dot{\dot{P}}$ and $\dot{\dot{Q}}$. Hence, to get the solution for $(\dot{P}, \dot{Q}, \dot{\dot{P}}, \dot{\dot{Q}})$, it requires to invert an $2n(n + 1) \times 2n(n + 1)$ matrix. If the rank condition (2.35) is observed, that is, using (W_1, U_1, Ω) to replace $(\dot{P}, \dot{Q})$, the solution would require to invert a matrix of dimension $n_* \times n_*$, where $n_* \triangleq n(n + n_c + 1) + n_c(n + 1)$. Thus, if $m << n$ and $l << n$, which is usually true, then $m_* << n_*$ and the algorithm discussed in this Appendix will be much more efficient. Furthermore, if (B.1)–(B.4) are first transformed to the bases in which A_P, A_Q, A_w and A_u are nearly diagonal, respectively, the cost of computation can be reduced significantly.

Next, we formulate a procedure to solve the derivative equations and correction equations in a desired basis. First, let T_P, T_Q, T_w and T_u be the transformation matrices such that $T_P^{-1} A_P T_P$, $T_Q^{-1} A_Q T_Q, T_w^{-1} A_w T_w$, and $T_u^{-1} A_u T_u$ are in suitable form according to the *basis* option described in Section 4.4. Next, make the replacements

$$A_P \leftarrow T_P^{-1} A_P T_P, \quad A_Q \leftarrow T_Q^{-1} A_Q T_Q, \quad A_w \leftarrow T_w^{-1} A_w T_w, \quad A_u \leftarrow T_u^{-1} A_u T_u,$$

$$\dot{P} \leftarrow T_P^{\mathrm{T}} \dot{P} T_P, \quad \dot{Q} \leftarrow T_Q^{-1} \dot{Q} T_Q^{-\mathrm{T}}, \quad \dot{\dot{P}} \leftarrow T_w^{\mathrm{T}} \dot{\dot{P}} T_w, \quad \dot{\dot{Q}} \leftarrow T_u^{-1} \dot{\dot{Q}} T_u^{-\mathrm{T}},$$

$$F_P \leftarrow T_P^{\mathrm{T}} F_P T_P, \quad F_Q \leftarrow T_Q^{-1} F_Q T_Q^{-\mathrm{T}}, \quad F_{\dot{P}} \leftarrow T_w^{\mathrm{T}} F_{\dot{P}} T_w, \quad F_{\dot{Q}} \leftarrow T_u^{-1} F_{\dot{Q}} T_u^{-\mathrm{T}},$$

$$W_1 \leftarrow T_u^{-1} W_1, \quad U_1 \leftarrow T_w^{\mathrm{T}} U_1, \quad W_2 \leftarrow T_u^{-1} W_2, \quad U_2 \leftarrow T_w^{\mathrm{T}} U_2.$$

Then we obtain

$$\phi_1(\dot{P}, \dot{Q}) \triangleq \mathrm{vec}\, U_1^{\mathrm{T}} T_w^{-1} T_u U_1^{\mathrm{T}} \alpha(\dot{P}, \dot{Q}) T_u^{\mathrm{T}} P B, \tag{B.35}$$

$$\phi_2(\dot{P}) \triangleq \mathrm{vec}\, B^{\mathrm{T}} T_P^{-\mathrm{T}} \dot{P}, \tag{B.36}$$

$$\phi_3(\dot{\dot{P}}) \triangleq \mathrm{vec}\, \dot{\dot{P}} T_w^{-1} T_u W_1 \Omega^{-1} W_1^{\mathrm{T}} T_u^{\mathrm{T}} P B, \tag{B.37}$$

$$\phi_4(\dot{\dot{P}}, \dot{\dot{Q}}) \triangleq \mathrm{vec}\, W_1^{\mathrm{T}} T_u^{\mathrm{T}} T_w^{-\mathrm{T}} \gamma(\dot{\dot{P}}, \dot{\dot{Q}}) T_w^{-1} Q C^{\mathrm{T}}, \tag{B.38}$$

$$\phi_5(\dot{Q}) \triangleq \mathrm{vec}\, C T_Q \dot{Q}, \tag{B.39}$$

$$\phi_6(\dot{\dot{Q}}) \triangleq \mathrm{vec}\, \dot{\dot{Q}} T_u^{\mathrm{T}} T_w^{-\mathrm{T}} U_1 \Omega^{-1} U_1^{\mathrm{T}} T_w^{-1} Q C^{\mathrm{T}}, \tag{B.40}$$

and

$$F_{a1} = -T_P^{\mathrm{T}} T_w^{-\mathrm{T}} U_1 \Omega, \quad G_{a1} = R_2^{-1} B^{\mathrm{T}} P T_u W_2 U_2^{\mathrm{T}} T_w^{-1} T_P, \quad F_{a2} = -T_P^{\mathrm{T}} T_w^{-\mathrm{T}} U_2 W_2^{\mathrm{T}} T_u^{\mathrm{T}} P B R_2^{-1},$$

$$G_{a2} = T_P^{-1} T_u W_1 U_1^{\mathrm{T}} T_w^{-1} T_P, \quad F_{a3} = -T_P^{\mathrm{T}} T_w^{-\mathrm{T}}, \quad G_{a3} = R_2^{-1} B^{\mathrm{T}} P T_u W_2 U_2^{\mathrm{T}} T_w^{-1} T_P,$$

$$F_{b1} = -T_Q^{-1} T_u W_1 \Omega, \quad G_{b1} = V_2^{-1} C Q T_w^{-\mathrm{T}} U_2 W_2^{\mathrm{T}} T_u^{\mathrm{T}} T_Q^{-\mathrm{T}}, \quad F_{b2} = -T_Q^{-1} T_u W_2 U_2^{\mathrm{T}} T_w^{-1} Q C^{\mathrm{T}} V_2^{-1},$$

$$G_{b2} = T_Q^{\mathrm{T}} T_w^{-\mathrm{T}} U_1 W_1^{\mathrm{T}} T_u^{\mathrm{T}} T_Q^{-\mathrm{T}}, \quad F_{b3} = -T_Q^{-1} T_u, \quad G_{b3} = V_2^{-1} C Q T_w^{-\mathrm{T}} U_2 W_2^{\mathrm{T}} T_u^{\mathrm{T}} T_Q^{-\mathrm{T}},$$

$$F_{c1} = U_1 \Omega, \quad G_{c1} = R_2^{-1} B^{\mathrm{T}} P T_u W_2 U_2^{\mathrm{T}}, \quad F_{c2} = -T_w^{\mathrm{T}} C^{\mathrm{T}} V_2^{-1}, \quad G_{c2} = T_Q^{\mathrm{T}} T_w^{-\mathrm{T}} U_1 \Omega U_1^{\mathrm{T}},$$

$$F_{c3} = U_2 W_2^{\mathrm{T}} T_u^{\mathrm{T}} P B R_2^{-1}, \quad G_{c3} = T_P^{-1} T_u W_1 U_1^{\mathrm{T}}, \quad F_{c4} = U_1 W_1^{\mathrm{T}} T_u^{\mathrm{T}} P B R_2^{-1}, \quad G_{c4} = T_P^{-1} T_w,$$

$$F_{d1} = W_1 \Omega, \quad G_{d1} = V_2^{-1} C Q T_w^{-\mathrm{T}} U_2 W_2^{\mathrm{T}}, \quad F_{d2} = W_2 U_2^{\mathrm{T}} T_w^{-1} Q C^{\mathrm{T}} V_2^{-1}, \quad G_{d2} = T_Q^{\mathrm{T}} T_w^{-\mathrm{T}} U_1 W_1^{\mathrm{T}},$$

$$F_{d3} = W_1 U_1^{\mathrm{T}} T_w^{-1} Q C^{\mathrm{T}} V_2^{-1}, \quad G_{d3} = T_Q^{\mathrm{T}} T_u^{-\mathrm{T}}, \quad F_{d4} = -T_u^{-1} B R_2^{-1}, \quad G_{d4} = T_P^{-1} T_u W_1 \Omega W_1^{\mathrm{T}},$$

satisfying (B.24)–(B.27) in the new basis.

Appendix C Conversion from $(\dot{Q}, \dot{P})$ to $(\dot{W}_1, \dot{U}_1, \dot{\Omega})$

Note that the following procedure is valid only in the original basis. It is desired to compute $\dot{W}_1$, $\dot{U}_1$ and $\dot{\Omega}$ satisfying (4.22)-(4.24). Note that (4.12) implies

$$U_1^{\mathrm{T}} W_1 = I_{n_m}, \quad U_1^{\mathrm{T}} W_2 = 0, \quad U_2^{\mathrm{T}} W_1 = 0. \tag{C.1}$$

Pre- and post-multiplying (4.22) by U and U^{T} respectively gives

$$\underline{\dot{Q}} = \underline{\dot{W}}_1 [\Omega \quad 0] + \begin{bmatrix} \dot{\Omega} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Omega \\ 0 \end{bmatrix} (\underline{\dot{W}}_1)^{\mathrm{T}}, \tag{C.2}$$

where

$$\underline{\dot{Q}} \triangleq U \dot{Q} U^{\mathrm{T}}, \tag{C.3}$$

$$\underline{\dot{W}}_1 \triangleq U \dot{W}_1 = \begin{bmatrix} U_1^{\mathrm{T}} \dot{W}_1 \\ U_2^{\mathrm{T}} \dot{W}_2 \end{bmatrix}. \tag{C.4}$$

Similarly, pre- and post-multiplying (4.23) by W^{T} and W respectively gives

$$\underline{\dot{P}} = \underline{\dot{U}}_1 [\Omega \quad 0] + \begin{bmatrix} \dot{\Omega} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Omega \\ 0 \end{bmatrix} \underline{\dot{U}}_1^{\mathrm{T}}, \tag{C.5}$$

where

$$\underline{\dot{P}} \triangleq W^{\mathrm{T}} \dot{P} W, \tag{C.6}$$

$$\underline{\dot{U}}_1 \triangleq W^{\mathrm{T}} \dot{U}_1 = \begin{bmatrix} W_1^{\mathrm{T}} \dot{U}_1 \\ W_2^{\mathrm{T}} \dot{U}_1 \end{bmatrix}. \tag{C.7}$$

Partition $\underline{\dot{W}}_1$ and $\underline{\dot{U}}_1$ as

$$\underline{\dot{W}}_1 = \begin{bmatrix} \underline{\dot{W}}_{11} \\ \underline{\dot{W}}_{21} \end{bmatrix}, \tag{C.8}$$

$$\underline{\dot{U}}_1 = \begin{bmatrix} \underline{\dot{U}}_{11} \\ \underline{\dot{U}}_{21} \end{bmatrix}. \tag{C.9}$$

It then follows from (C.4) and (C.7)–(C.9) that (4.24) is equivalent to

$$\underline{\dot{W}}_{11} = -\underline{\dot{U}}_{11}^{\mathrm{T}}. \tag{C.10}$$

It now follows from (C.8) that (C.2) is equivalent to

$$\begin{bmatrix} \underline{\dot{Q}}_{11} & (\underline{\dot{Q}}_{21})^{\mathrm{T}} \\ \underline{\dot{Q}}_{21} & 0 \end{bmatrix} = \begin{bmatrix} \underline{\dot{W}}_{11}\Omega & 0 \\ \underline{\dot{W}}_{21}\Omega & 0 \end{bmatrix} + \begin{bmatrix} \dot{\Omega} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Omega(\underline{\dot{W}}_{11})^{\mathrm{T}} & \Omega(\underline{\dot{W}}_{21})^{\mathrm{T}} \\ 0 & 0 \end{bmatrix}, \tag{C.11}$$

and from (C.9) that (C.5) is equivalent to

$$\begin{bmatrix} \underline{\dot{P}}_{11} & (\underline{\dot{P}}_{21})^{\mathrm{T}} \\ \underline{\dot{P}}_{21} & 0 \end{bmatrix} = \begin{bmatrix} \underline{\dot{U}}_{11}\Omega & 0 \\ \underline{\dot{U}}_{21}\Omega & 0 \end{bmatrix} + \begin{bmatrix} \dot{\Omega} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} \Omega(\underline{\dot{U}}_{11})^{\mathrm{T}} & \Omega(\underline{\dot{U}}_{21})^{\mathrm{T}} \\ 0 & 0 \end{bmatrix}. \tag{C.12}$$

Furthermore, (C.11) is equivalent to

$$\underline{\dot{Q}}_{11} = \underline{\dot{W}}_{11}\Omega + \dot{\Omega} + \Omega(\underline{\dot{W}}_{11})^{\mathrm{T}}, \tag{C.13}$$

$$\underline{\dot{Q}}_{21} = \underline{\dot{W}}_{21}\Omega. \tag{C.14}$$

Similarly, equation (C.12) is equivalent to

$$\underline{\dot{P}}_{11} = \underline{\dot{U}}_{11}\Omega + \dot{\Omega} + \Omega(\underline{\dot{U}}_{11})^{\mathrm{T}}, \tag{C.15}$$

$$\underline{\dot{P}}_{21} = \underline{\dot{U}}_{21}\Omega. \tag{C.16}$$

31

Now, (C.14) and (C.16) imply respectively that

$$\dot{\underline{W}}_{21} = \dot{\underline{Q}}_{21}\Omega^{-1}, \qquad (C.17)$$

$$\dot{\underline{U}}_{21} = \dot{\underline{P}}_{21}\Omega^{-1}. \qquad (C.18)$$

Furthermore, substituting (C.10) into (C.13) yields

$$\dot{\underline{Q}}_{11} = -(\dot{\underline{U}}_{11})^{\mathrm{T}}\Omega + \dot{\Omega} - \Omega\dot{\underline{U}}_{11}. \qquad (C.19)$$

Denote the (i, j) elements of $\dot{\underline{P}}_{11}, \dot{\underline{Q}}_{11}$, and $\dot{\underline{U}}_{11}$ respectively by $\dot{\underline{p}}_{ij}, \dot{\underline{q}}_{ij}$, and $\dot{\underline{u}}_{ij}$. Then we can rewrite (C.15) and (C.19) as

$$\dot{\underline{p}}_{ij} = \dot{\underline{u}}_{ij}\omega_j + \delta_{ij}\dot{\omega}_i + \omega_i\dot{\underline{u}}_{ji}, \quad i, j \in \{1, 2, \ldots, n_m\}, \qquad (C.20)$$

$$\dot{\underline{q}}_{ij} = -\dot{\underline{u}}_{ji}\omega_j + \delta_{ij}\dot{\omega}_i - \omega_i\dot{\underline{u}}_{ij}, \quad i, j \in \{1, 2, \ldots, n_m\}, \qquad (C.21)$$

where

$$\delta_{ij} \triangleq \begin{cases} 1 \text{ for } i = j \\ 0 \text{ for } i \neq j. \end{cases} \qquad (C.22)$$

Next, assume $i = j$. Then subtracting (C.21) from (C.20) gives

$$\dot{\underline{u}}_{ii} = \frac{\dot{\underline{p}}_{ii} - \dot{\underline{q}}_{ii}}{4\omega_i}. \qquad (C.23)$$

Now, assume $i \neq j$. Multiplying (C.20) by (ω_j/ω_i) and adding the resultant equations to (C.21) gives

$$\dot{\underline{u}}_{ij} = \frac{\omega_j\dot{\underline{p}}_{ij} + \omega_i\dot{\underline{q}}_{ij}}{\omega_j^2 - \omega_i^2}, \quad \omega_i \neq \omega_j, \qquad (C.24a)$$

or, if $\omega_i = \omega_j$,

$$\dot{\underline{u}}_{ij} = \frac{\dot{\underline{p}}_{ij} - \dot{\underline{q}}_{ji}}{4\omega_i}. \qquad (C.24b)$$

Now, $\dot{\underline{U}}_{11}$ is defined by (C.23) and (C.24) and $\dot{\underline{U}}_{21}$ by (C.18). $\dot{\underline{W}}_{11}$ is then defined by (C.10) and $\dot{\underline{W}}_{21}$ by (C.17). $\dot{\underline{W}}_1$ and $\dot{\underline{U}}_1$ are now defined respectively by (C.8) and (C.9). Using (4.24) it follows from (C.4) and (C.7) that $\dot{W}_1$ and $\dot{U}_1$ are given respectively by

$$\dot{W}_1 = W\dot{\underline{W}}_1, \qquad (C.25)$$

$$\dot{U}_1 = U^{\mathrm{T}}\dot{\underline{U}}_1. \qquad (C.26)$$

From (C.22) it follows that

$$\dot{\omega}_i = \frac{\dot{\underline{p}}_{ii} + \dot{\underline{q}}_{ii}}{2}, \qquad (C.27)$$

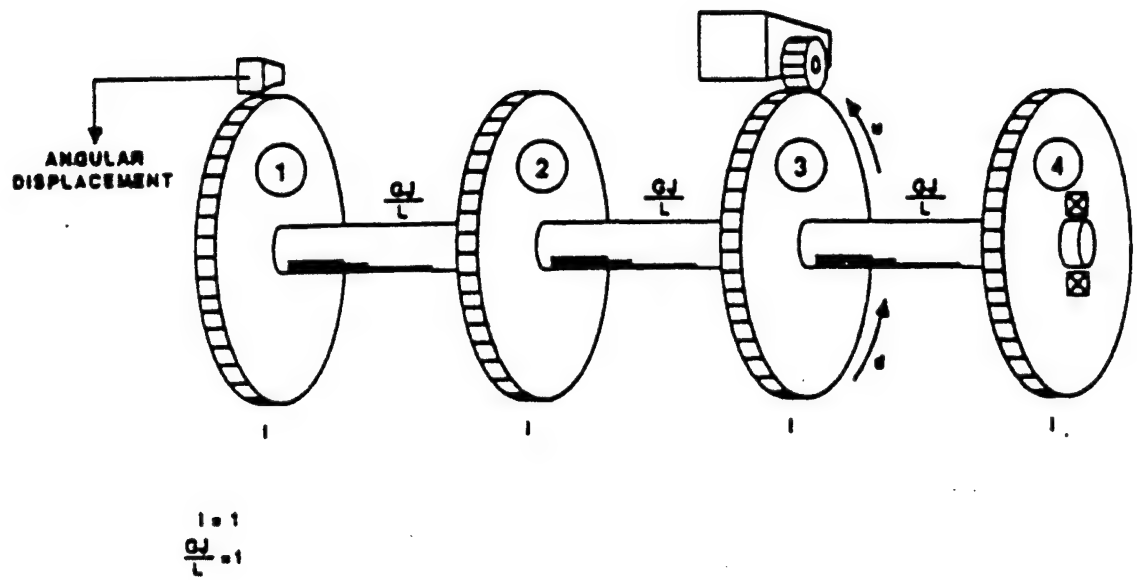which defines $\dot{\Omega}$.

32

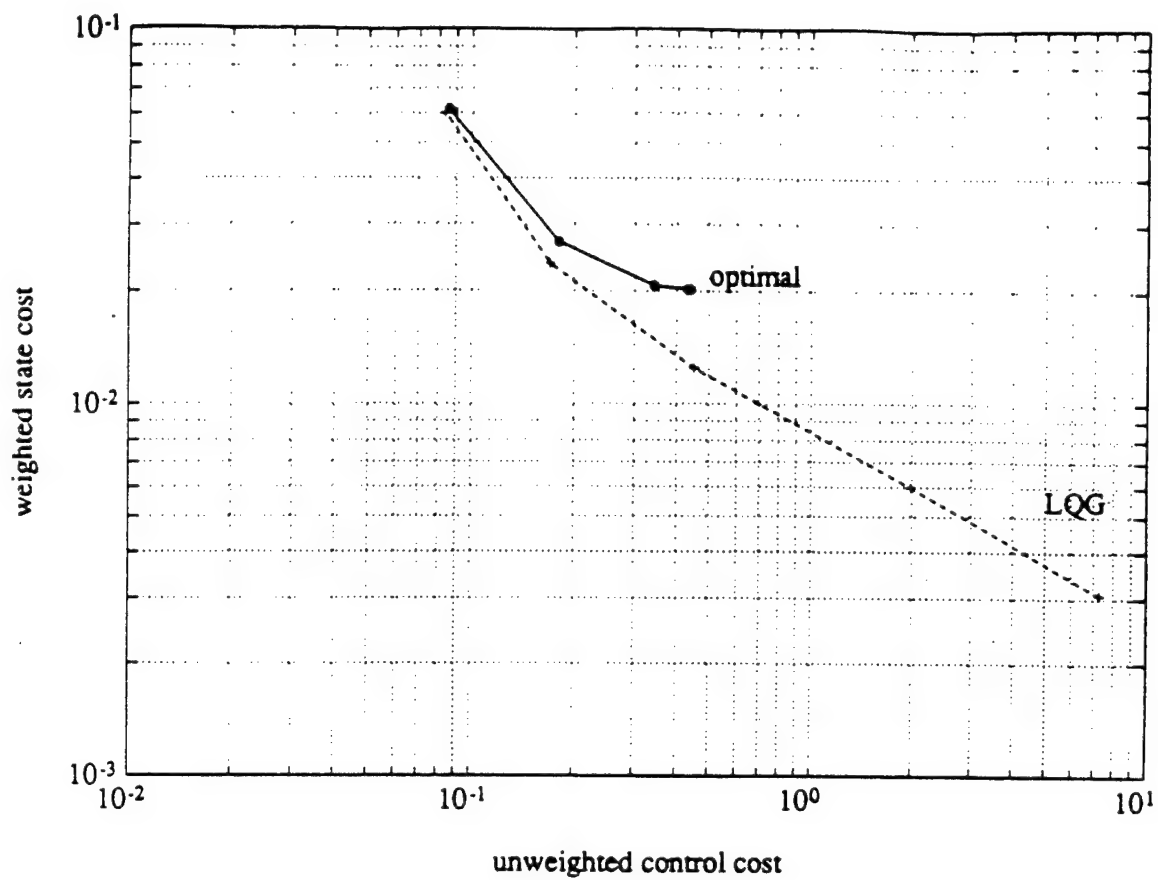Figure 5.1. The Four Disk Model

Figure 5.2. Performance Curves for the 2nd-order Controllers for $q_2 \in (10, 1, \ldots, 10^{-4})$
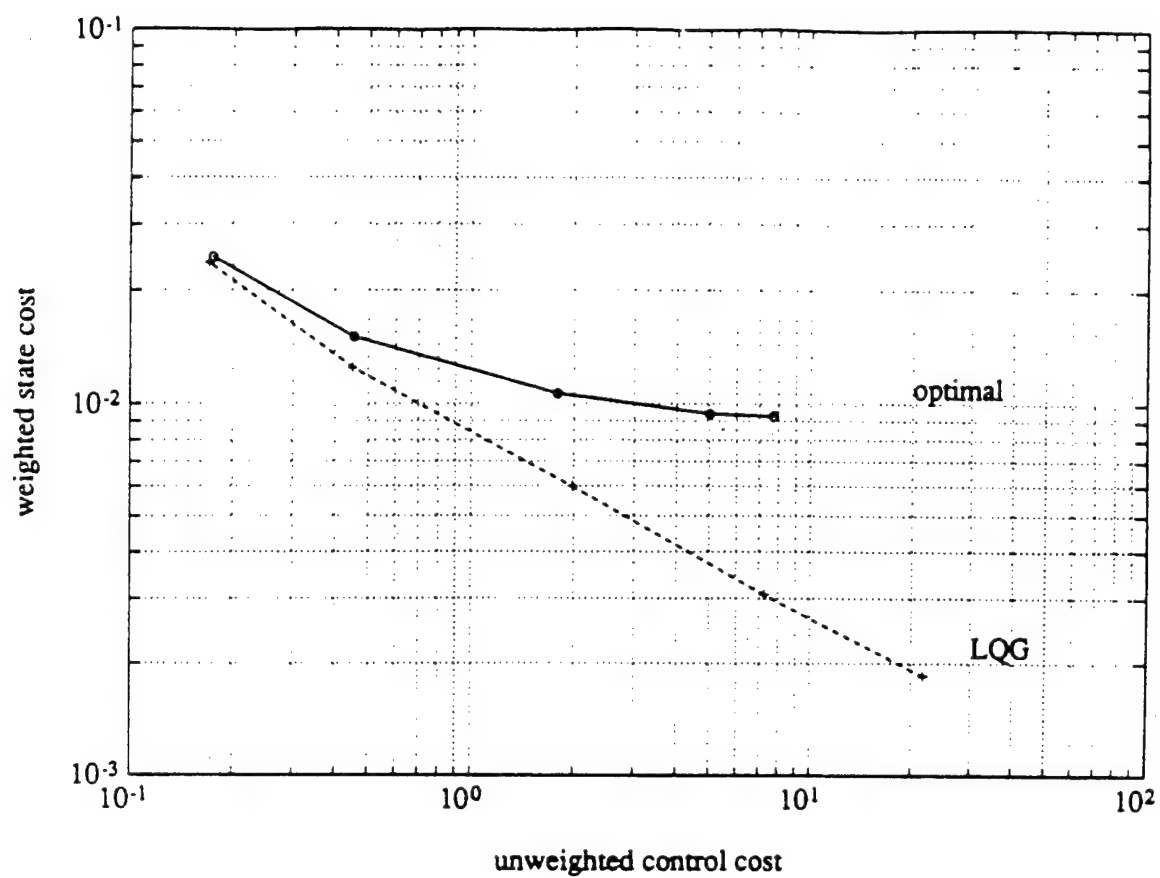
Figure 5.3. Performance Curves for the 4th-order Controllers for $q_2 \in (1, 0.1, \ldots, 10^{-5})$
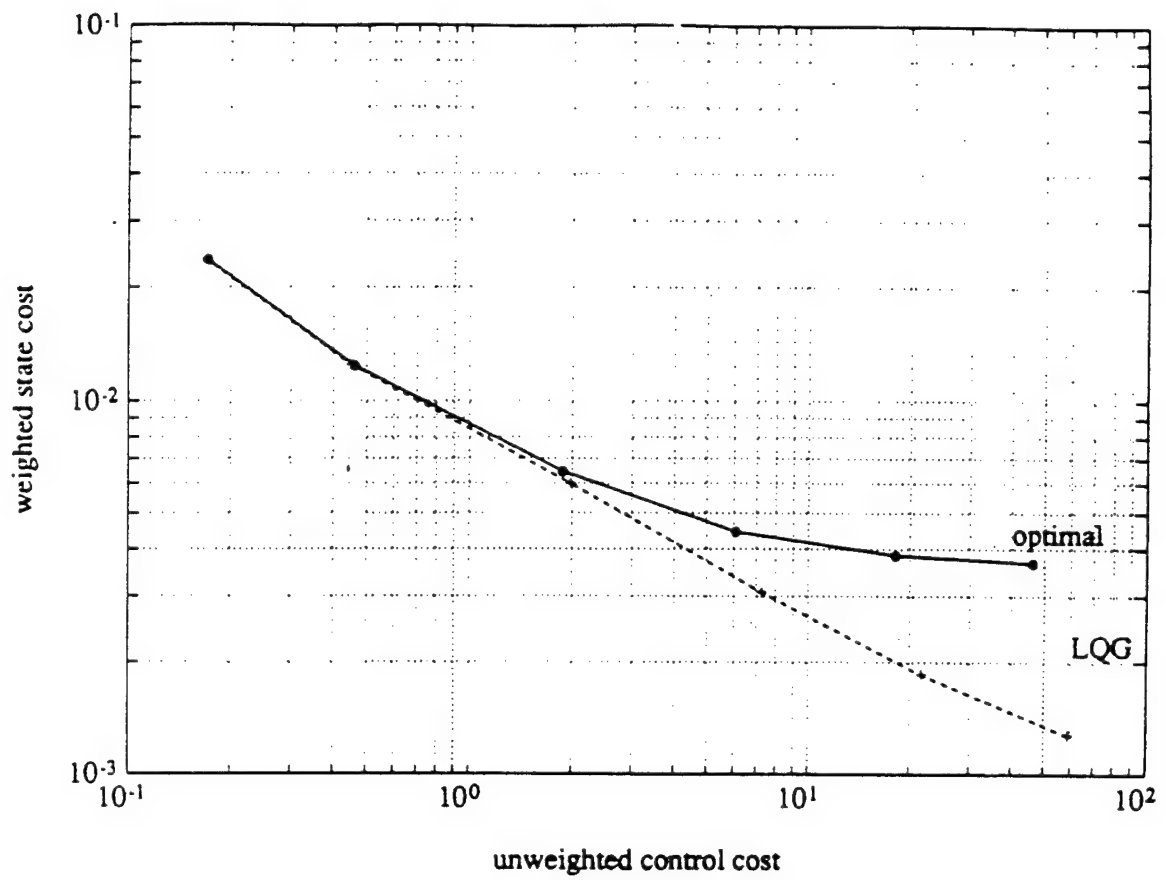
Figure 5.4. Performance Curves for the 6th-order Controllers for $q_2 \in (1, 0.1, \ldots, 10^{-6})$
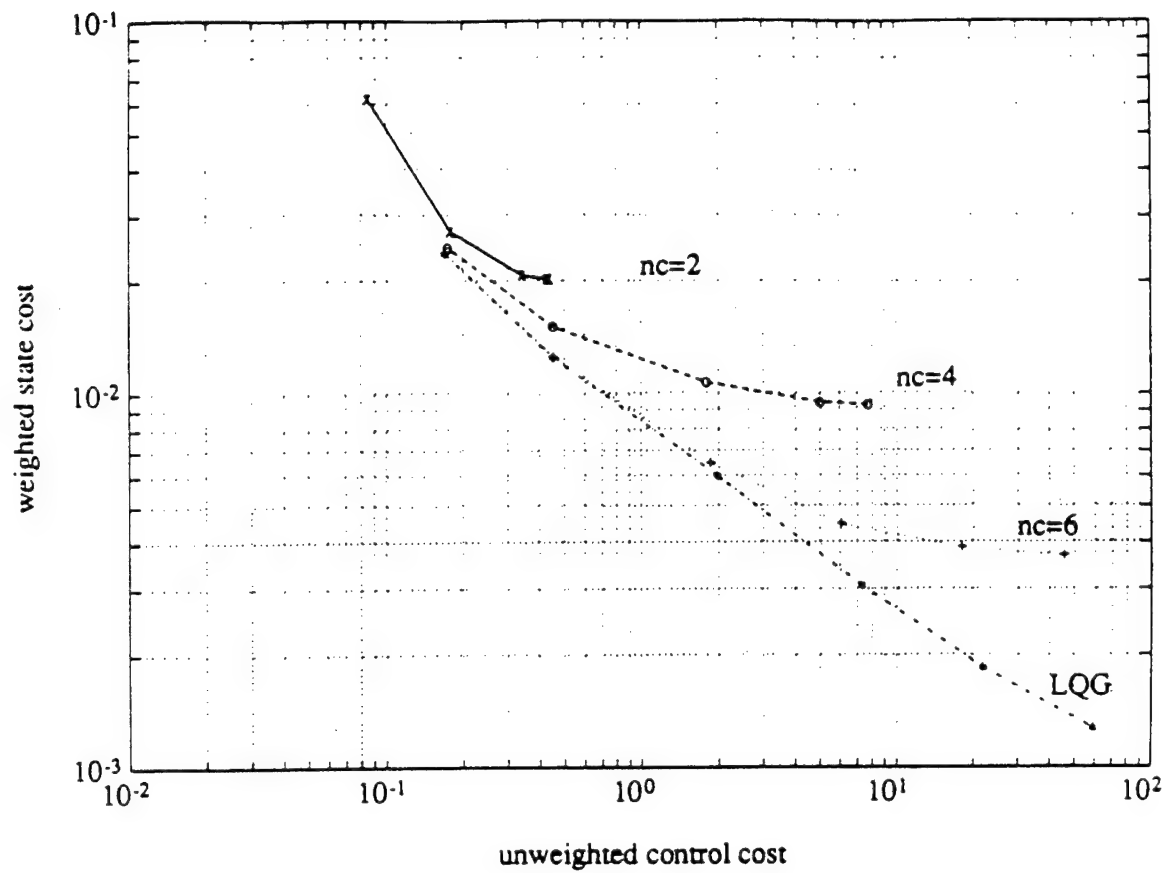
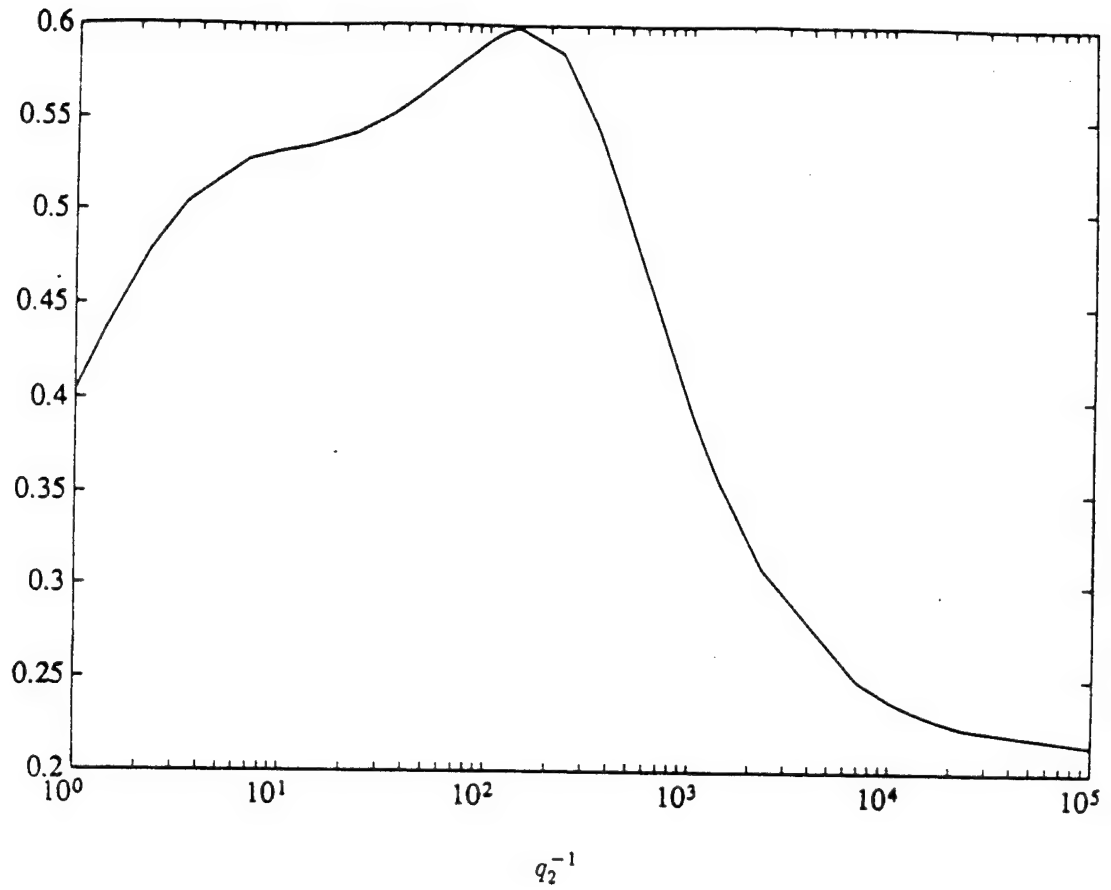Figure 5.5. Comparison of the Performance Curves for Various Order Controllers for Four Disk Example

Figure 5.6. $\|P\|_F$ as a Function of Control Authority (q_2^{-1}) for Four Disk Example with $n_c = 4$
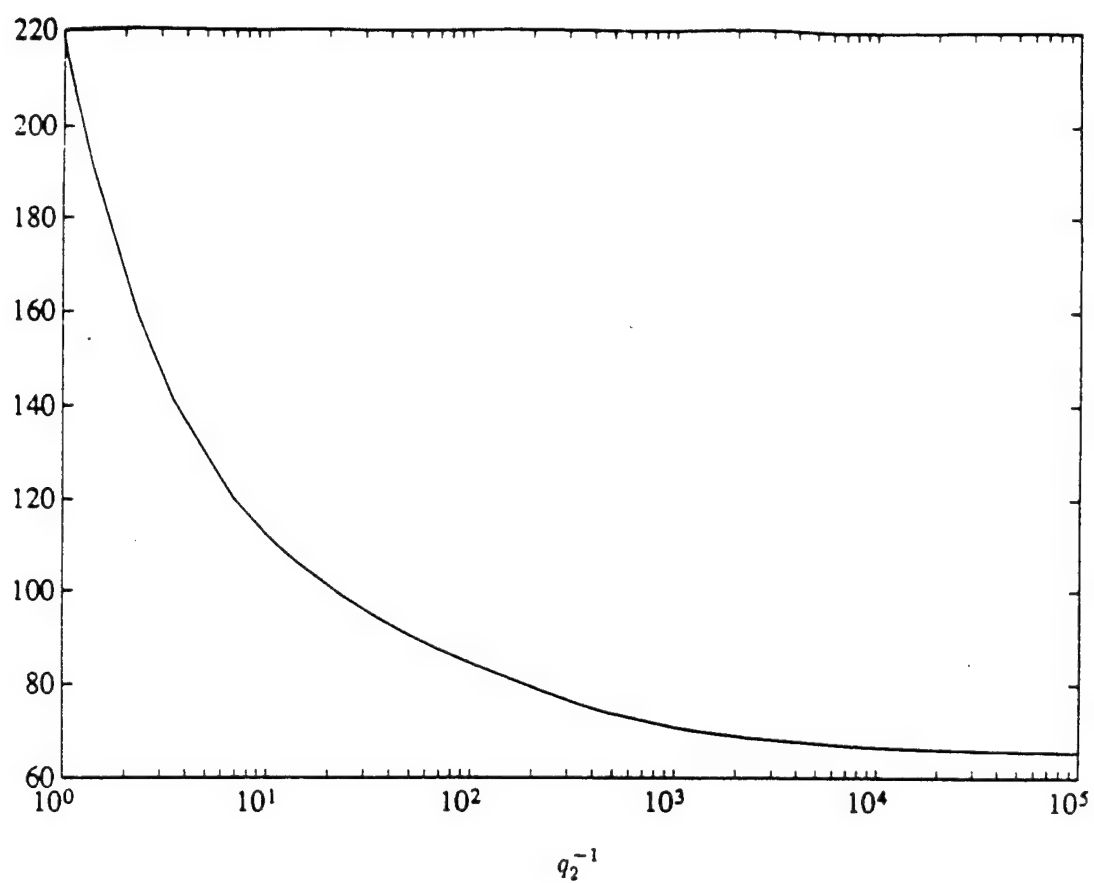
Figure 5.7. $\|Q\|_F$ as a Function of Control Authority (q_2^{-1}) for Four Disk Example with $n_c = 4$
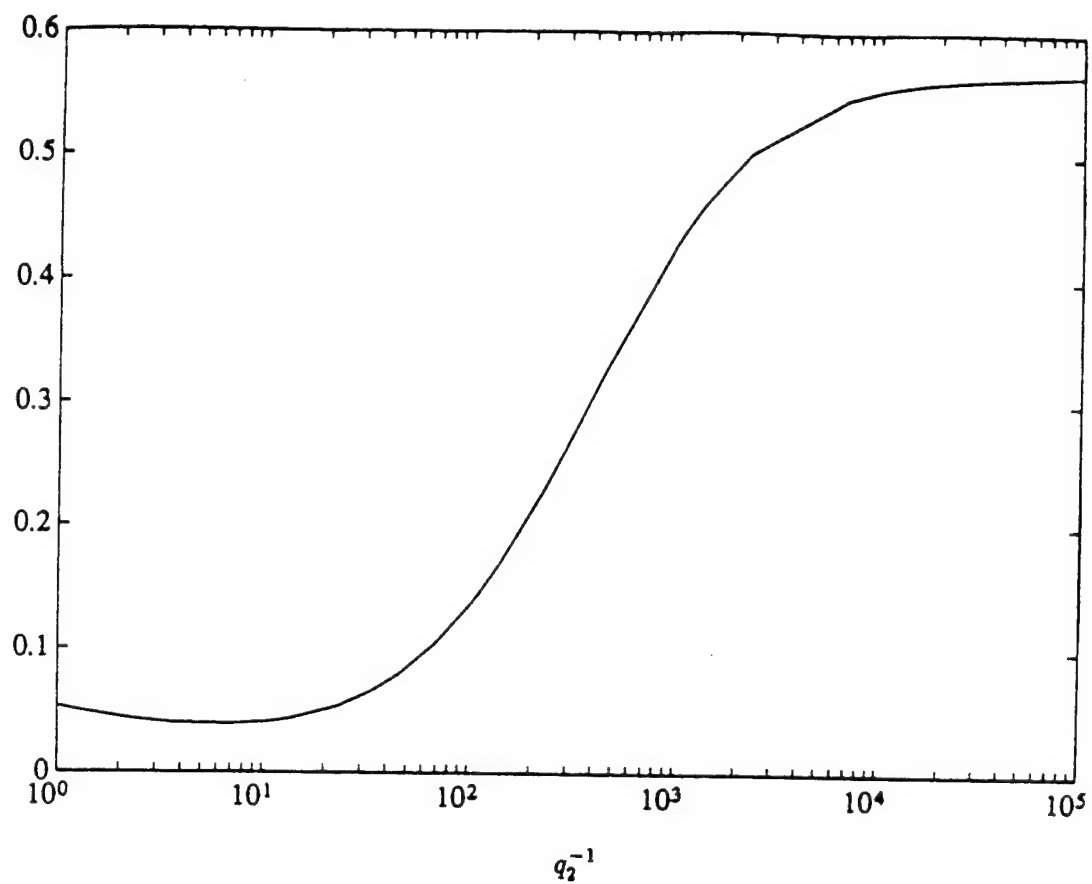
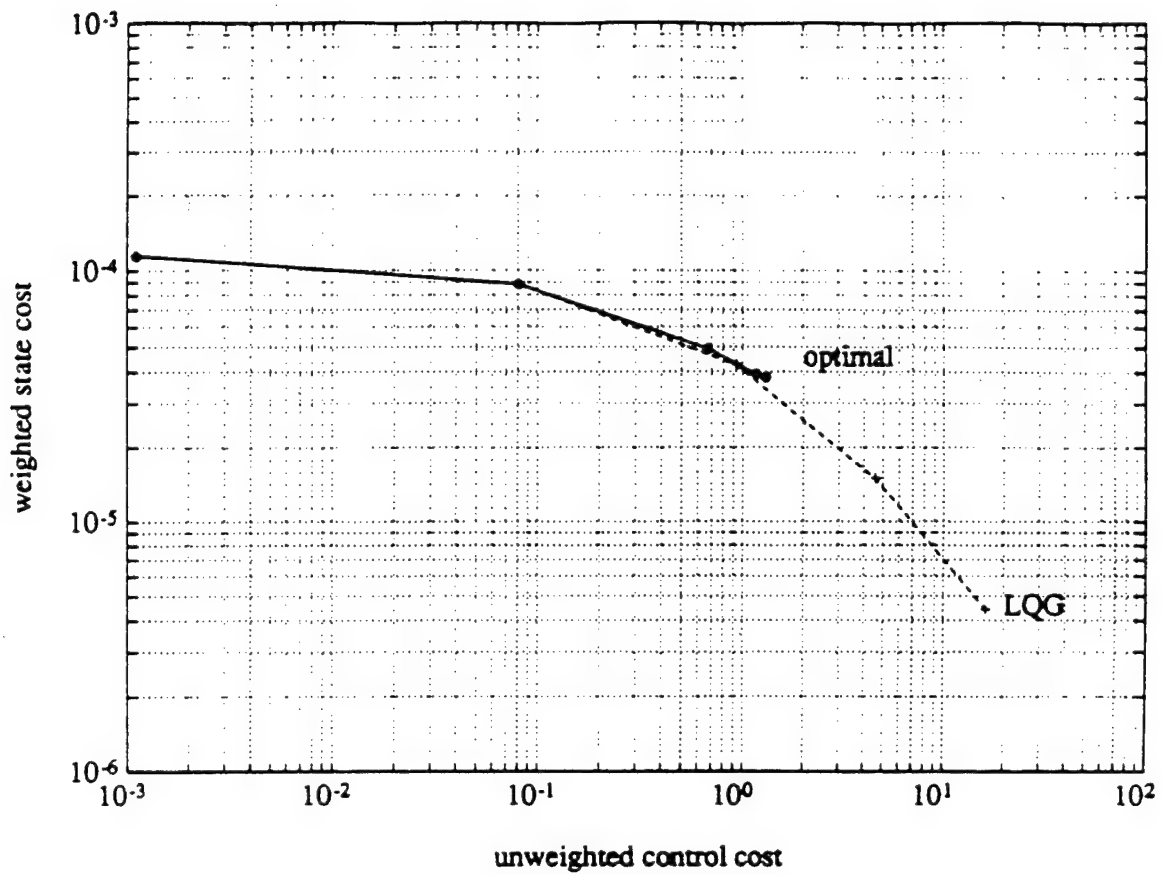Figure 5.8. $\|\hat{P}\|_F$ as a Function of Control Authority (q_2^{-1}) for Four Disk Example with $n_c = 4$
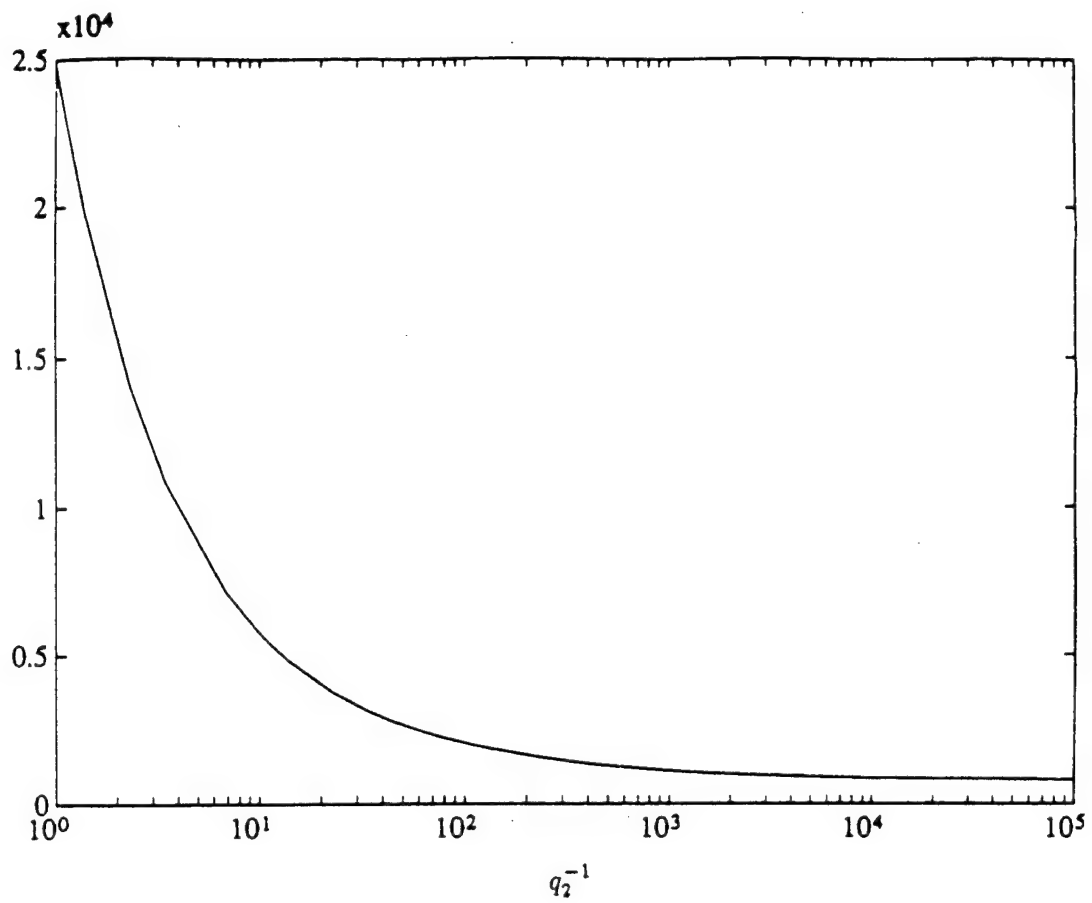
Figure 5.9. $\|\hat{Q}\|_F$ as a Function of Control Authority (q_2^{-1}) for Four Disk Example with $n_c = 4$

Figure 5.10.   Performance Curves for the 8th-order Controllers for ACES Structure.

References

[1] Makila, P. M., Toivonen, H. T., "Computational Methods for Parametric LQ Problems–a Survey," *IEEE Transactions on Automatic Control*, Vol. 32, 1987, pp. 658-671.

[2] Levine, W. S., Johnson, T. L., and Athans, M.,"Optimal Limited State Variable Feedback Controllers for Linear Systems," *IEEE Transactions on Automatic Control*, Vol. 16, 1971, pp. 785-793.

[3] Martin, G. D., and Bryson, A. E., Jr., "Attitude Control of a Flexible Spacecraft," *AIAA Journal of Guidance, Control, and Dynamics*, Vol. 3, 1980, pp. 37–41.

[4] Mukhopadhyay, V., Newsom, Journal R., and Abel, I., "Reduced-Order Optimal Control Law Synthesis for Flutter Suppression," *Journal of Guidance, Control and Dynamics*, Vol. 4, 1982, pp. 389-395.

[5] Ly, U-L, Bryson, A. E., and Cannon, R. H., "Design of Low-Order Compensators Using Parameter Optimization," *Automatica*, Vol. 21, 1985, pp. 315-318.

[6] Mukhopadhyay, V., "Stability Robustness Improvement Using Constrained Optimization Techniques," *AIAA Journal of Guidance Control and Dynamics*, Vol. 10, 1987, pp. 172-177.

[7] Mukhopadhyay, V., "Digital Robust Control Law Synthesis Using Constrained Optimization," *Journal of Guidance, Control and Dynamics*, Vol. 12, 1989, pp. 175-181.

[8] Anderson, B. D. O., "Second-Order Convergent Algorithms for the Steady-State Riccati-Equation," *Int. J. Control*, Vol. 28, 1978, pp. 295–306.

[9] Makila P. M., "Parametric LQ Control," *Int. J. Control*, Vol. 41, 1985, pp. 1413–1428.

[10] Moerder, D. D., and Calise, A. J., "Convergence of a Numerical Algorithm for Calculating Optimal Output Feedback Gains," *IEEE Transactions on Automatic Control*, Vol. 30, 1985, pp. 900–903.

[11] Toivonen, H.T., and Makila, P. M., "Newton's Method for Solving Parametric Linear Quadratic Control Problems," *Int. J. Control*, Vol. 46, 1987, pp. 897–911.

[12] Peterson, L. D., "Optimal Projection Control of an Experimental Truss Structure," *Journal of Guidance, Control and Dynamcis*, Vol. 14, No. 2, 1991, pp. 241–250.

[13] Mercadal, M., "Homotopy Approach to Optimal, Linear Quadratic, Fixed Architecture Compensation," *Journal of Guidance, Control and Dynamics*, Vol. 14, 1991, pp. 1224–1233.

[14] Collins, E. G., Jr., Phillips, D. J., and Hyland, D. C., "Robust Decentralized Control Laws for the ACES Structure," *Control Systems Magazine*, 1991, pp. 62–70.

[15] Collins, E. G., Jr., Davis, L.D., and Richter, S., "Design of Reduced-Order, H_2 Optimal Controllers Using a Homotopy Algorithm ," *Int. J. Control*, 1993, to appear.

[16] Kuhn, U. and Schmidt, G., "Fresh Look into the Design and Computation of Optimal Output Feedback Controls for Linear Multivariable Systems," *International Journal of Control*, Vol. 46, 1987. pp. 75-95.

[17] Ge, Y., Collins, E. G., Jr., Watson, L. T., and Davis, L. D., "A Input Normal Form

Homotopy for the L^2 Optimal Model Order Reduction Problem," 1993, submitted to *Int. J. Control.*

[18] Ge, Y., Collins, E. G., Jr., Watson, L. T., and Davis, L. D., "A Comparison of Homotopies for Alternative Formulations of the L^2 Optimal Model Order Reduction Problem," 1993, submitted to *Int. J. Control.*

[19] Richter, S. and Collins, E. G. Jr., "A Homotopy Algorithm for Reduced-Order Controller Design using the Optimal Projection Equations," *Proc. IEEE Conference on Decision and Control,* 1989, pp. 932–933.

[20] Hyland, D. C. and Bernstein, D. S., "The Optimal Projection Equations for Fixed-order Dynamic Compensation," *IEEE Transactions on Automatic Control,* Vol. 29, 1984, pp. 1034–1037.

[21] Richter, S., Davis, L. D., and Collins, E. G., Jr., "Efficient Computation of the Solutions to Modified Lyapunov Equations," *SIAM Journal of Matrix Analysis and Applications,* 1993, pp. 420–431.

[22] Zigić, D., Watson, L. T., Collins, E. G., Jr., and Bernstein, D. S., "Homotopy Methods for Solving the Optimal Projection Equations for the H_2 Reduced Order Model Problem," *International Journal of Control* , Vol. 56, 1992, pp. 173–191.

[23] Zigić, D., Watson, L. T., Collins, E. G., Jr., and Bernstein, D. S., "Homotopy Approaches to the H_2 Reduced Order Model Problem," *Journal of Mathematical Systems, Estimation, Control,* 1993, to appear.

[24] Collins, E. G. Jr., Ying, S. S., Haddad, W. M., and Richter, S.,"An Efficient, Numerically Robust Homotopy Algorithm for H_2 Model Reduction Using the Optimal Projection Equations," 1993, Submitted to *International J. Control.*

[25] Collins, E. G., Jr., Davis, L.D., and Richter, S., "A Homotopy Algorithm for Maximum Entropy Design," *Proc. Amer. Contr. Conf.,* 1993, San Francisco, CA, pp. 1010–1019.

[26] Zigić, D., Watson, L. T., and Collins, E. G., Jr., "A Homotopy Method for Solving Riccati Equations on a Shared Memory Parallel Computer," *Sixth SIAM Conference on Parallel Processing for Scientific Computing* , 1993, pp. 614–617.

[27] Rao, C.R., and Mitra, S.K., *Generalized Inverse of Matrices and its Applications,* 1971, John Wiley and Sons, New York.

[28] Bernstein, D.S. and Haddad, W.M., "Robust Stability and Performance via Fixed-Order Dynamic Compensation with Guaranteed Cost Bounds," *Math. Control Signal Systems,* Vol. 3, 1990, pp. 139–163.

[29] Lloyd, N. G., *Degree Theory,* 1978, Cambridge University Press, London.

[30] Avila, J. H., "The Feasibility of Continuation Methods for Nonlinear Equations," *SIAM Journal of Numerical Analysis,* Vol. 11, 1974, pp. 104–144.

[31] Wacker, H., *Continuation Methods,* 1978, Academic Press, New York.

[32] Alexander, J. C., and Yorke, J. A., "The homotopy Continuation Method: Numerically Implementable Topological Procedures," *Transactions of the American Mathematical Society,*

Vol. 242, 1978, pp. 271–284.

[33] Garcia, C. B., and Zangwill, W. I., *Pathways to Solutions, Fixed Points and Equilibria*, 1981, Prentice-Hall, Englewood Cliffs, NJ.

[34] Eaves, B. C., Gould, F. J., Peoitgen, J. O., and Todd, M. J., *Homotopy Methods and Global Convergence*, 1983, Plenum Press, New York.

[35] Watson, L. T., "Numerical Linear Algebra Aspects of Globally Convergent Homotopy Methods," *SIAM Review*, Vol. 28, 1986, pp. 529–545.

[36] Richter, S. L., and DeCarlo, R. A., "Continuation Methods: Theory and Applications," *IEEE Transactions on Circuits and Systems*, Vol. 30, 1983, pp. 347–352.

[37] Richter, S., and DeCarlo, R., "A Homotopy Method for Eigenvalue Assignment Using Decentralized State Feedback," *IEEE Transactions on Automatic Control*, Vol. 29, No. 2, 1984, pp. 148–158.

[38] Turner, J. D., and Chun, H. M., "Optimal Distributed Control of a Flexible Spacecraft During a Large-Angle Maneuver," *Journal of Guidance, Control and Dynamics*, Vol. 7, 1984, pp. 257–264.

[39] Dunyak, J. P., Junkins, J. L., and Watson, L. T., "Robust Nonlinear Least Squares Estimation Using the Chow-Yorke Homotopy Method," *Journal of Guidance, Control and Dynamics*, Vol. 7, 1984, pp. 752–755.

[40] Lefebvre, S., Richter, S., and DeCarlo, R., " A Continuation Algorithm for Eigenvalue Assignment by Decentralized Constant-Output Feedback," *International Journal of Control*, Vol. 41, 1985, pp. 1273–1292.

[41] Sebok, D. R., Richter, S. and DeCarlo, R., "Feedback Gain Optimization in Decentralized Eigenvalue Assignment," *Automatica*, Vol. 22, 1986, pp. 433-447.

[42] Horta, L. G., Juang, J.-N., and Junkins, J. L., "A Sequential Linear Optimization Approach for Controller Design," *Journal of Guidance, Control and Dynamics*, 9, 1986, pp. 699–703.

[43] Kabamba, P. T., Longman, R. W., and Jian-Guo, S., 1987, "A Homotopy Approach to the Feedback Stabilization of Linear Systems," *Journal of Guidance, Control and Dynamics*, Vol. 10, 1987, pp. 422–432.

[44] Shin, Y. S., Haftka, R. T., Watson, L. T., and Plaut, R. H., "Tracking Structural Optima as a Function of Available Resources by a Homotopy Method," *Computer Methods in Applied Mechanics and Engineering*, Vol. 70, 1988, pp. 151–164.

[45] Rakowska, J., Haftka, R. T., and Watson, L. T., "Tracing the Efficient Curve for Multi-Objective Control-Structure Optimization," *Comput. Systems. Engrg.*, Vol. 2, 1991, pp. 461–472.

[46] Watson, L. T., "ALGORITHM 652 HOMPACK: A Suite of Codes for Globally Convergent Homotopy Algorithms," *ACM Transactions on Mathematical Software*, Vol. 13, 1987, pp. 281–310.

[47] Collins, E. G. Jr., Haddad, W. M., and Ying, S. S., "Construction of Low Authority,

Nearly Non-Minimal LQG Compensators for Reduced-Order Control Design," 1993, Submitted to *IEEE Transactions on Automatic Control and the 1994 American Control Conference.*

[48] Yousuff, A., and Skelton, R.E., "A Note on Balanced Controller Reduction," *IEEE Transactions on Automatic Control,* Vol. AC-29, 1984, pp. 254–257.

[49] De Villemagne, C., and Skelton, R. E., "Controller Reduction Using Canonical Interactions," *IEEE Transactions on Automatic Control,* Vol. AC-33, 1988, pp. 740–750.

[50] Fletcher, R., *Practical Methods of Optimization,* John Wiley and Sons, New York, 1987.

[51] Brewer, J.W., "Kronecker Products and Matrix Calculus in System Theory," *IEEE Trans. Circuit and Systems,* Vol. 25, 1978, pp. 772–781.

[52] Cannon, R. H., and Rosenthal, D. E., "Experiments in Control of Flexible Structures with Noncolocated Sensors and Actuators," *AIAA Journal of Guidance, Control and Dynamics,* Vol. 7, 1984, pp. 546-553.

[53] Anderson, B. D. O. and Liu, Y., " Controller Reduction: Concepts and Approaches," *IEEE Transactions on Automatic Control,* Vol. 34, 1989, pp. 802–812.

[54] Liu, Y., Anderson, B. D. O. and Ly, U-L., " Coprime Factorization Controller Reduction with Bezout Identity Induced Frequency Weighting,," *Automatica,* Vol. 26, 1990, pp. 233-249.

[55] Hyland, D. C. and Richter, S., "On Direct Versus Indirect Methods for Reduced-order Controller Design," *IEEE Transactions on Automatic Control,* Vol. 35, 1990, pp. 377-379.

[56] Irwin, R. D., Jones, V. L., Rice, S. A., Seltzer, S. M., and Tollison, D. J., *Active Control Technique Evaluation for Spacecraft (ACES),* Final Report to Flight Dynamics Lab of Wright Aeronautical Labs, Report No., AFWAL-TR-88-3038, June 1988.

Appendix L:

Computation of the Complex Structured Singular Value Using Fixed Structure Dynamic D-Scales

Appendix M:

New Frequency Domain Performance Bounds
for Structural Systems with Actuater and Sensor Dynamics

April 1994

New Frequency Domain Performance Bounds for Uncertain Structural Systems with Actuator and Sensor Dynamics

by

Wassim M. Haddad
School of Aerospace Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332
(404) 894-1078
FAX: (404) 894-2760
wm.haddad@aerospace.gatech.edu

Emmanuel G. Collins, Jr.
Harris Corporation
Government Aerospace Systems
Division
Melbourne, Florida 32902
(407) 727-6358
FAX: (407) 727-4016
ecollins@x102a.ess.harris.com

David C. Hyland
Harris Corporation
Government Aerospace Systems
Division
Melbourne, Florida 32902
(407) 729-2138
FAX: (407) 727-4016

Vijaya-Sekhar Chellaboina
School of Aerospace Engineering
Georgia Institute of Technology
Atlanta, Georgia 30332
(404) 894-3000
FAX: (404) 894-2760
gt3582a@acme.gatech.edu

Abstract

A new majorant robustness analysis test that yields frequency dependent performance bounds for closed-loop uncertain vibrational systems with frequency, damping, and mode shape uncertainties is developed. Specifically, for closed-loop systems consisting of uncertain positive real plants in series with sensor and actuator dynamics and controlled by strictly positive real compensators, performance bounds are developed by decomposing the equivalent compensator (which includes the actuator and sensor dynamics) to a positive real part and a non-positive real part and using concepts of M-matrices and majorant analysis.

Key Words: Frequency domain performance bounds, robust stability and performance, majorant bounds, uncertain vibrational systems

Running Title: Frequency domain performance bounds

---

1. Introduction

The analysis of uncertain dynamical systems for robust stability and performance remains one of the most important issues in modern feedback control theory. This necessitates the development of efficient analysis tools that allow a control system to be analyzed for robustness with respect to structured and unstructured uncertainty in the design model. Hence, considerable effort has been devoted to robust analysis in the recent years. Many of the developments in robust analysis have focused exclusively on stability robustness while ignoring robust performance. However, it is well known that robust performance is of paramount importance in practice. Specifically, even though stability robustness addresses the qualitative question as to whether or not a system remains stable for all plant perturbations within a specified class of uncertainties it is important to *quantitatively* investigate the performance degradation within the region of robust stability. In practice it is often desirable to determine the *worst-case* performance as a measure of degradation. The interested reader is referred to Bernstein and Haddad (1990) and the references therein for a more complete exposition of the robust stability and performance analysis problem.

In a recent paper by Hyland *et al.* (1994) the tools of majorant analysis used to develop robust stability and performance tests in Hyland and Bernstein (1987), Collins and Hyland (1989), and Hyland and Collins (1989), (1991) were extended to positive real plants controlled by strictly positive real compensators. Specifically, using the logarithmic norm in the context of majorant analysis, *new* majorant robustness analysis tests were developed that yield frequency dependent performance bounds for frequency, damping, and mode shape uncertainty in positive real vibrational systems. For this class of systems the positive real majorant bounds developed in Hyland *et al.* (1994) yield much less conservative robustness (stability and performance) predictions over previous norm based majorant performance bounds (Hyland and Collins, 1989) and the performance bound resulting from complex structured singular value analysis (Hyland *et al.*, 1994; Packard and Doyle, 1993).

The main purpose of this paper is to extend the results presented in Hyland *et al.*

1

(1994) to uncertain positive real structural systems in series with actuator and sensor dynamics. It is well known that in this case the resulting system is no longer positive real and hence the results of Hyland *et al.* (1994) can no longer be applied. Using the framework developed in Hyland *et al.* (1994) we develop new frequency domain performance bounds for this more general class of uncertain structural systems. Specifically, the results are developed by decomposing the equivalent compensator consisting of the original strictly positive real compensator along with the actuator and sensor dynamics into a positive real part and a non-positive real part and using the concepts of M-matrices and majorant analysis. To demonstrate the effectiveness of the proposed approach we apply our results to an Euler-Bernoulli beam with closely spaced frequency uncertainty and actuator and sensor dynamics.

Notation

In the following notation, the matrices and vectors are in general assumed to be complex.

$\mathbb{R}$	set of real numbers
$\mathbb{C}$	set of complex numbers
I_p	$p \times p$ identity matrix
Z^{H}	complex conjugate transpose of matrix Z
z_{ij} or Z_{ij}	(i, j) element of matrix Z
$\mathrm{diag}\{z_1, \ldots, z_n\}$	diagonal matrix with listed diagonal elements
$Y \leq\leq Z$	$y_{ij} \leq z_{ij}$ for each i and j, where Y and Z are real matrices with identical dimensions
$|\alpha|$	absolute value of complex scalar α
$\det(Z)$	determinant of square matrix Z
$\|x\|_2$	Euclidean norm of vector $x \; (= \sqrt{x^{\mathrm{H}}x})$
$\sigma_{\min}(Z), \sigma_{\max}(Z)$	minimum, maximum singular values of matrix Z
$\|Z\|_{\mathrm{F}}$	Frobenius norm of matrix $Z \big(= (\mathrm{tr}\, ZZ^{\mathrm{H}})^{\frac{1}{2}}\big)$
$\rho(Z)$	spectral radius of a square matrix Z
$\lambda_{\min}(Z), \lambda_{\max}(Z)$	minimum, maximum eigenvalues of the Hermitian matrix Z
$\max\{Y_1, \ldots, Y_n\}$	$= \overline{Y}$ where $\overline{y}_{ij} = \max\{y_{1,ij}, y_{2,ij}, \ldots, y_{n,ij}\}$
$\mathcal{L}[z(t)]$	Laplace transform of $z(t)$

2. Mathematical Preliminaries

In this section we establish several definitions and two key lemmas. A *nonnegative matrix* Z is a matrix with nonnegative elements, i.e, $Z \geq\geq 0$. A *block-norm matrix*

2

(Ostrowski, 1975) is a nonnegative matrix each of whose elements is the norm of the corresponding subblock of a given partitioned matrix. The *modulus matrix* of $A \in \mathbb{C}^{m \times n}$ is the $m \times n$ nonnegative matrix

$$|A|_M \triangleq [|a_{ij}|]. \tag{2.1}$$

Note that the modulus matrix is a special case of a block norm matrix. Let $B \in \mathbb{C}^{n \times p}$. Subsequent analysis will use the relation

$$|AB|_M \leq\leq |A|_M |B|_M. \tag{2.2}$$

A *majorant* (Dahlquist, 1983) is an element-by-element upper bound for a modulus matrix (or, more generally, a block norm matrix). Specifically, $\hat{A}$ is an $m \times n$ majorant of $A \in \mathbb{C}^{m \times n}$ if

$$|A|_M \leq\leq \hat{A}. \tag{2.3}$$

Let $Z \in \mathbb{C}^{n \times n}$. Then $\check{Z} \in \mathbb{R}^{n \times n}$ is an $n \times n$ *minorant* (Dahlquist, 1983) of Z if

$$\check{z}_{ii} \leq |z_{ii}|, \tag{2.4a}$$

$$\check{z}_{ij} \leq -|z_{ij}|, \quad i \neq j. \tag{2.4b}$$

The following lemma is a direct consequence of the above definitions.

Lemma 2.1. Let Z_d and Z_{od} denote, respectively, the diagonal and off-diagonal components of $Z \in \mathbb{C}^{n \times n}$, such that

$$Z_d = \text{diag}\{z_{ii}\}_{i=1}^n, \quad Z_{od} = Z - Z_d. \tag{2.5}$$

Then, if $\check{Z}_d$ is an $n \times n$ minorant of Z_d and $\hat{Z}_{od}$ is a majorant of Z_{od}, $\check{Z}_d - \hat{Z}_{od}$ is a minorant of Z.

A matrix $P \in \mathbb{R}^{n \times n}$ is an *M-matrix* (Fiedler and Ptak, 1962; Seneta, 1973; and Berman and Plemmons, 1979) if it has nonpositive off-diagonal elements (i.e., $p_{ij} \leq 0$ for $i \neq j$) and positive principal minors. Recall that the inverse of an M-matrix is a nonnegative matrix (Fiedler and Ptak, 1962; Seneta, 1973; and Berman and Plemmons, 1979).

3

Lemma 2.2.(Dahlquist, 1983). Assume $Z \in \mathbb{C}^{n \times n}$ and let $\check{Z}$ be an $n \times n$ minorant of Z. If in addition $\check{Z}$ is an M-matrix, then Z is nonsingular and

$$|Z^{-1}|_M \leq\leq \check{Z}^{-1}. \tag{2.6}$$

3. Robust Stability and Performance for Uncertain Vibrational Systems with Actuator and Sensor Dynamics

We begin by considering the following n^{th}-order, uncertain, matrix second-order vibrational system with proportional damping and rate measurements:

$$\ddot{\eta}(t) + 2\Lambda\Omega\dot{\eta}(t) + \Omega^2\eta(t) = Bu(t) + Dw(t), \tag{3.1a}$$

$$y(t) = C\dot{\eta}(t), \tag{3.1b}$$

$$z(t) = E\dot{\eta}(t), \tag{3.1c}$$

where

$$\Omega = \text{diag}\{\Omega_i\}_{i=1}^n, \quad \Omega_i > 0, \; i = 1, 2, \ldots, n, \tag{3.2}$$

$$\Lambda = \text{diag}\{\zeta_i\}_{i=1}^n, \quad \zeta_i > 0, \; i = 1, 2, \ldots, n, \tag{3.3}$$

$u \in \mathbb{R}^{n_u}$ is the control vector, $w \in \mathbb{R}^{n_w}$ is the disturbance variable or reference signal, $y \in \mathbb{R}^{n_y}$ represents the rate measurements, and $z \in \mathbb{R}^{n_z}$ represents the performance variables (restricted to be linear functions of the modal rates). It is assumed that

$$\Omega \in \boldsymbol{\Omega} \triangleq \{\Omega_0 + \Delta\Omega : |\Delta\Omega|_M \leq\leq \widehat{\Delta\Omega}\}, \tag{3.4}$$

$$\Lambda \in \boldsymbol{\Lambda} \triangleq \{\Lambda_0 + \Delta\Lambda : |\Delta\Lambda|_M \leq\leq \widehat{\Delta\Lambda}\}, \tag{3.5}$$

$$B \in \mathbf{B} \triangleq \{B_0 + \Delta B : |\Delta B|_M \leq\leq \widehat{\Delta B}\}, \tag{3.6}$$

$$D \in \mathbf{D} \triangleq \{D_0 + \Delta D : |\Delta D|_M \leq\leq \widehat{\Delta D}\}, \tag{3.7}$$

$$C \in \mathbf{C} \triangleq \{C_0 + \Delta C : |\Delta C|_M \leq\leq \widehat{\Delta C}\}, \tag{3.8}$$

$$E \in \mathbf{E} \triangleq \{E_0 + \Delta E : |\Delta E|_M \leq\leq \widehat{\Delta E}\}. \tag{3.9}$$

Next, define

$$H_1 \triangleq (\Omega, \Lambda), \tag{3.10}$$

$$H_2 \triangleq (B, C), \tag{3.11}$$

$$H_3 \triangleq (D, E), \tag{3.12}$$

4

and define $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_3$ to be the corresponding uncertainty sets, i.e.,

$$\mathbf{H}_1 \triangleq \{(\Omega, \Lambda) : \Omega \in \boldsymbol{\Omega}, \ \Lambda \in \boldsymbol{\Lambda}\}, \tag{3.13}$$

$$\mathbf{H}_2 \triangleq \{(B, C) : B \in \mathbf{B}, \ C \in \mathbf{C}\}, \tag{3.14}$$

$$\mathbf{H}_3 \triangleq \{(D, E) : D \in \mathbf{D}, \ E \in \mathbf{E}\}. \tag{3.15}$$

Additionally, define

$$\mathbf{H} \triangleq \mathbf{H}_1 \cup \mathbf{H}_2 \cup \mathbf{H}_3. \tag{3.16}$$

Note that $\mathbf{H}_1$ is the uncertainty set corresponding to errors in the frequencies and damping ratios while $\mathbf{H}_2$ and $\mathbf{H}_3$ are uncertainty sets corresponding to errors in the mode shapes. It follows from (3.4)–(3.9) that $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_3$ are arcwise connnected.

Furthermore, let

$$\theta(s) \triangleq \mathcal{L}[\dot{\eta}(t)], \tag{3.17}$$

so that (3.1) has the s-domain representation

$$\Phi^{-1}(H_1, s)\theta(H, s) = Bu(s) + Dw(s), \tag{3.18a}$$

$$y(H, s) = C\theta(s), \tag{3.18b}$$

$$z(H, s) = E\theta(s), \tag{3.18c}$$

where

$$\Phi(H_1, s) \triangleq \operatorname{diag}\{\phi_i(H_1, s)\}_{i=1}^{n}, \tag{3.19}$$

and

$$\phi_i(H_1, s) \triangleq \frac{s}{s^2 + 2\zeta_i \Omega_i s + \Omega_i^2}. \tag{3.20}$$

Note that for all $H_1 \in \mathbf{H}_1$, $\Phi(H_1, s)$ is strictly positive real, so that

$$\Phi(H_1, j\omega) + \Phi^{\mathrm{H}}(H_1, j\omega) > 0, \quad H_1 \in \mathbf{H}_1, \quad \omega \in (0, \infty). \tag{3.21}$$

If, alternatively, the system is undamped, that is, $\zeta_i = 0$, $i = 1, \ldots, n$, then (3.19) is positive real.

5

To make the model more realistic we now include sensor and actuator dynamics that are assumed to be known. These dynamics could be empirically determined via hardware experimentation. The matrix transfer function of actuator dynamics (Ψ_a) and the matrix transfer function of sensor dynamics (Ψ_s) are given by

$$\Psi_a(s) \triangleq \text{diag}\{\Psi_{a,i}(s)\}_{i=1}^{n_a}, \tag{3.22}$$

$$\Psi_s(s) \triangleq \text{diag}\{\Psi_{s,i}(s)\}_{i=1}^{n_y}. \tag{3.23}$$

Appending these dynamics to the system (3.18) yields

$$\Phi^{-1}(H_1, s)\theta(H, s) = B\Psi_a(s)u(s) + Dw(s), \tag{3.24a}$$

$$y(H, s) = \Psi_s(s)C\theta(s), \tag{3.24b}$$

$$z(H, s) = E\theta(s). \tag{3.24c}$$

Next, assume that the linear feedback law

$$u(s) = -K(s)y(s) \tag{3.25}$$

stabilizes the nominal system, i.e., the system (3.24) with $H_1 = (\Omega_0, \Lambda_0)$ and $H_2 = (B_0, C_0)$. Furthermore, assume colocated velocity feedback so that $B = C^T$. Substituting (3.25) into (3.24a) gives

$$[\Phi^{-1}(H_1, s) + F(H_2, s)]\theta(H, s) = Dw(s), \tag{3.26}$$

where

$$F(H_2, s) \triangleq B\Psi(s)B^T, \tag{3.27}$$

$$\Psi(s) \triangleq \Psi_a(s)K(s)\Psi_s(s).$$

Now, define $F_{\text{pr}}(H_2, j\omega)$ to be the positive real part and $F_{\text{npr}}(H_2, j\omega)$ to be the non-positive real part of the "equivalent" compensator $F(H_2, j\omega)$, respectively, so that

$$F(H_2, j\omega) = F_{\text{pr}}(H_2, j\omega) + F_{\text{npr}}(H_2, j\omega), \qquad \omega \in (0, \infty), \tag{3.28}$$

6

where

$$F_{\mathrm{pr}}(H_2, j\omega) \triangleq \begin{cases} F(H_2, j\omega), & \lambda_{\min}(\Psi(j\omega) + \Psi^{\mathrm{H}}(j\omega)) \geq 0, \\ 0, & \text{otherwise}, \end{cases} \tag{3.29}$$

and

$$F_{\mathrm{npr}}(H_2, j\omega) \triangleq F(H_2, j\omega) - F_{\mathrm{pr}}(H_2, j\omega), \tag{3.30}$$

for all $\omega \in (0, \infty)$. Similarly, define

$$\Psi_{\mathrm{pr}}(j\omega) \triangleq \begin{cases} \Psi(j\omega), & \lambda_{\min}(\Psi(j\omega) + \Psi^{\mathrm{H}}(j\omega)) \geq 0, \\ 0, & \text{otherwise}, \end{cases}$$

and

$$\Psi_{\mathrm{npr}}(j\omega) \triangleq \Psi(j\omega) - \Psi_{\mathrm{pr}}(j\omega), \qquad \omega \in (0, \infty).$$

The following three lemmas are key to the development of the robust stability and performance bounds presented in this paper.

Lemma 3.1. If $\Phi(H_1, s)$ is strictly postive real for all $H_1 \in \mathbf{H}$ and $F_{\mathrm{pr}}(H_2, j\omega)$ is given by (3.29) then $\left[\Phi^{-1}(H_1, j\omega) + \Phi^{-\mathrm{H}}(H_1, j\omega)\right] + \left[F_{\mathrm{pr}}(H_2, j\omega) + F_{\mathrm{pr}}^{\mathrm{H}}(H_2, j\omega)\right] > 0$ and hence

$$\det[\Phi^{-1}(H_1, j\omega) + F_{\mathrm{pr}}(H_2, j\omega)] \neq 0, \quad \omega \in (0, \infty). \tag{3.31}$$

Proof. First we show that $\Phi(H_1, j\omega)$ is invertible and $\Phi^{-1}(H_1, j\omega)$ is strictly positive real. Let $x \in \mathbb{C}^n$, $x \neq 0$, and $\lambda \in \mathbb{C}$ be such that $\Phi(H_1, j\omega)x = \lambda x$ and hence $x^{\mathrm{H}}\Phi^{\mathrm{H}}(H_1, j\omega) = \lambda^{\mathrm{H}}x^{\mathrm{H}}$. Then $x^{\mathrm{H}}\left[\Phi(H_1, j\omega) + \Phi^{\mathrm{H}}(H_1, j\omega)\right]x > 0$ implies that $\mathrm{Re}\,\lambda > 0$. Hence $\det\Phi(H_1, j\omega) \neq 0$. Now note that

$$\Phi^{-1}(H_1, j\omega) + \Phi^{-\mathrm{H}}(H_1, j\omega) = \Phi^{-1}(H_1, j\omega)\left[\Phi(H_1, j\omega) + \Phi^{\mathrm{H}}(H_1, j\omega)\right]\Phi^{-\mathrm{H}}(H_1, j\omega) > 0,$$

which implies that $\Phi^{-1}(H_1, s)$ is strictly positive real. Next, since for all $H_2 \in \mathbf{H}_2$, $F_{\mathrm{pr}}(H_2, j\omega) + F_{\mathrm{pr}}^{\mathrm{H}}(H_2, j\omega) \geq 0$ it follows that

$$\Phi^{-1}(H_1, j\omega) + \Phi^{-\mathrm{H}}(H_1, j\omega) + F_{\mathrm{pr}}(H_2, j\omega) + F_{\mathrm{pr}}^{\mathrm{H}}(H_2, j\omega) > 0,$$

for all $H_1 \in \mathbf{H}_1$ and $H_2 \in \mathbf{H}_2$. Now (3.31) is immediate. $\qquad \square$

For simplicity of exposition we define

$$\Gamma(H, \jmath\omega) \triangleq [\Phi^{-1}(H_1, \jmath\omega) + F_{\text{pr}}(H_2, \jmath\omega)]^{-1}, \qquad \omega \in (0, \infty). \qquad (3.32)$$

Furthermore, define $\mathcal{S} : \text{IR} \to \text{IR}$ as

$$\mathcal{S}(\alpha) \triangleq \begin{cases} \alpha, & \alpha \geq 0 \\ 0, & \alpha < 0. \end{cases}$$

The next two lemmas are a direct consequence of Theorems 4.1 and 4.2 and Theorems 6.1 and 6.2 of Hyland et $al.$ (1994). The proofs follow from majorant analysis and standard singular value inequalities and hence are omitted.

Lemma 3.2. (Hyland et $al.$, 1994). If $\Phi(H_1, s)$ is given by (3.19) and $F_{\text{pr}}(H_2, \jmath\omega)$ is given by (3.29) then

$$\max_{H \in \mathbf{H}} |\Gamma(H, \jmath\omega)|_{\text{M}} \leq\leq \hat{\Gamma}_0(\jmath\omega),$$

where

$$\hat{\Gamma}_0(\jmath\omega) = \underline{p}^{-1}(\jmath\omega)U_n, \qquad (3.33)$$

$$\underline{p}(\jmath\omega) = \max\Bigg\{ \min_k 2(\zeta_{0,k} - \widehat{\Delta\zeta_k})(\Omega_{0,k} - \widehat{\Delta\Omega_k})$$
$$+ \frac{1}{2}\bigg[\mathcal{S}\big(\sigma_{\min}(B_0 M(\jmath\omega)) - \sigma_{\max}(M(\jmath\omega))\|\widehat{\Delta B}\|_{\text{F}}\big)\bigg]^2,$$
$$\min_k \big(\frac{1}{\omega}(\Omega_{0,k} - \widehat{\Delta\Omega_k})^2 - \omega\big) - \frac{1}{2}\sigma_{\max}(\Psi_{\text{pr}}(\jmath\omega) - \Psi_{\text{pr}}^{\text{H}}(\jmath\omega))(\sigma_{\max}(B_0) + \|\widehat{\Delta B}\|_{\text{F}})^2,$$
$$\min_k \big(\omega - \frac{1}{\omega}(\Omega_{0,k} + \widehat{\Delta\Omega_k})^2\big) - \frac{1}{2}\sigma_{\max}(\Psi_{\text{pr}}(\jmath\omega) - \Psi_{\text{pr}}^{\text{H}}(\jmath\omega))(\sigma_{\max}(B_0) + \|\widehat{\Delta B}\|_{\text{F}})^2\Bigg\},$$
$$(3.34)$$

$$\Psi_{\text{pr}}(\jmath\omega) + \Psi_{\text{pr}}^{\text{H}}(\jmath\omega) = M(\jmath\omega)M^{\text{H}}(\jmath\omega),$$

and U_n denotes the $n \times n$ matrix with all unity elements.

Lemma 3.3. (Hyland et $al.$, 1994). Assume $\Phi(H_1, s)$ is given by (3.19) and $F_{\text{pr}}(H_2, \jmath\omega)$ is given by (3.29) and let $F_{\text{d}}(H_2, \jmath\omega)$ and $F_{\text{od}}(H_2, \jmath\omega)$, respectively, denote the diagonal and off-diagonal matrices corresponding to $F_{\text{pr}}(H_2, \jmath\omega)$, such that

$$F_{\text{d}}(H_2, s) \triangleq \text{diag}\big\{f_{\text{pr},ii}(H_2, s)\big\}_{i=1}^n, \qquad (3.35)$$

$$F_{\text{od}}(H_2, s) \triangleq F_{\text{pr}}(H_2, s) - F_{\text{d}}(H_2, s). \qquad (3.36)$$

Let $\Pi(\jmath\omega)$ be given by

$$\Pi(\jmath\omega) = P(\jmath\omega) - \hat{F}_{\mathrm{od}}(\jmath\omega), \tag{3.37}$$

where $P(\jmath\omega)$ satisfies

$$P(\jmath\omega) = \mathrm{diag}\{p_{kk}(\jmath\omega)\}_{k=1}^{n}, \tag{3.38}$$

and

$$
\begin{aligned}
p_{kk}(\jmath\omega) = \max\bigg\{ & 2(\zeta_{0,k} - \widehat{\Delta\zeta_k})(\Omega_{0,k} - \widehat{\Delta\Omega_k}) \\
& + \frac{1}{2}\lambda_{\min}(\Psi_{\mathrm{pr}}(\jmath\omega) + \Psi_{\mathrm{pr}}^{\mathrm{H}}(\jmath\omega))\sum_{l=1}^{m}\big[\mathcal{S}(B_{0,kl} - \widehat{\Delta B_{kl}})\big]^2, \\
& \min_{\Omega\in\Omega}\left|\frac{\Omega_{0,k}^2}{\omega} - \omega\right| - \frac{1}{2}\sigma_{\max}(\Psi_{\mathrm{pr}}(\jmath\omega) - \Psi_{\mathrm{pr}}^{\mathrm{H}}(\jmath\omega))\sum_{l=1}^{m}\big[|B_{0,kl}| + \widehat{\Delta B_{kl}}\big]^2\bigg\}.
\end{aligned}
\tag{3.39}
$$

In addition, let $\hat{F}_{\mathrm{od}}(\jmath\omega)$ satisfying (3.37) be given by

$$[\hat{F}_{\mathrm{od}}(\jmath\omega)]_{ij} = \sigma_{\max}(\Psi_{\mathrm{pr}}(\jmath\omega))\Big[\sum_{k=1}^{m}(|B_{0,ik}| + \widehat{\Delta B_{ik}})^2\Big]^{\frac{1}{2}}\Big[\sum_{k=1}^{m}(|B_{0,jk}| + \widehat{\Delta B_{jk}})^2\Big]^{\frac{1}{2}}, \quad i \neq j. \tag{3.40}$$

Then, if $\Pi(\jmath\omega)$ is an M-matrix,

$$\max_{H\in\mathbf{H}}|\Gamma(H,\jmath\omega)|_{\mathrm{M}} \leq\leq \hat{\Gamma}_1(\jmath\omega),$$

where

$$\hat{\Gamma}_1(\jmath\omega) = \Pi^{-1}(\jmath\omega).$$

Next, define $\hat{\Gamma}(\jmath\omega)$ and $\hat{F}_{\mathrm{npr}}(\jmath\omega)$ such that

$$[\hat{\Gamma}(\jmath\omega)]_{ij} \triangleq \min([\hat{\Gamma}_0(\jmath\omega)]_{ij}, [\hat{\Gamma}_1(\jmath\omega)]_{ij}), \quad \omega \in (0,\infty), \tag{3.41}$$

and

$$\max_{H_2\in\mathbf{H}_2}|F_{\mathrm{npr}}(H_2,\jmath\omega)|_{\mathrm{M}} \leq\leq \hat{F}_{\mathrm{npr}}(\jmath\omega). \tag{3.42}$$

Now, note that $F_{\mathrm{npr}}(\jmath\omega)$ is given by

$$[F_{\mathrm{npr}}(\jmath\omega)]_{ij} = \sum_{l=1}^{m}\sum_{k=1}^{m} b_{il}\Psi_{\mathrm{npr},lk}b_{jk},$$

9

and $|[F_{\mathrm{npr}}(\jmath\omega)]_{ij}| \leq [\hat{F}_{\mathrm{npr}}(\jmath\omega)]_{ij}$ where $[\hat{F}_{\mathrm{npr}}(\jmath\omega)]_{ij}$ is given by

$$[\hat{F}_{\mathrm{npr}}(\jmath\omega)]_{ij} = \sigma_{\max}(\Psi_{\mathrm{npr}}(\jmath\omega)) \Big[\sum_{k=1}^{m}(|B_{0,ik}| + \widehat{\Delta B}_{ik})^2\Big]^{\frac{1}{2}} \Big[\sum_{k=1}^{m}(|B_{0,jk}| + \widehat{\Delta B}_{jk})^2\Big]^{\frac{1}{2}}.$$

Next we present the main result of this paper which gives robust stability and performance bounds for the uncertain vibrational system described by (3.24) and (3.25).

Theorem 3.1. The dynamic system given by (3.24) and (3.25) is asymptotically stable for all $H \in \mathbf{H}$, if

$$\rho\big(\hat{\Gamma}(\jmath\omega)\hat{F}_{\mathrm{npr}}(\jmath\omega)\big) < 1, \quad \omega \in (0,\infty). \tag{3.43}$$

Furthermore, the output $z(H, \jmath\omega)$ satisfies the bound

$$\max_{H \in \mathbf{H}} |z(H, \jmath\omega)|_{\mathrm{M}} \leq\leq |E|_{\mathrm{M}}\big[I_n - \hat{\Gamma}(\jmath\omega)\hat{F}_{\mathrm{npr}}(\jmath\omega)\big]^{-1}\hat{\Gamma}(\jmath\omega)|Dw(\jmath\omega)|_{\mathrm{M}}, \quad \omega \in (0,\infty). \tag{3.44}$$

Proof. It follows from the multivariable Nyquist criterion that in order to establish asymptotic stability of the closed-loop uncertain system given by (3.24) and (3.25) it suffices to show that $\det[\Phi^{-1}(H_1, \jmath\omega) + F(H_2, \jmath\omega)] \neq 0$ for all $\omega \in (0,\infty)$ and $H \in \mathbf{H}$. Using the definition of a minorant it follows that $I - \hat{\Gamma}(\jmath\omega)\hat{F}_{\mathrm{npr}}(\jmath\omega)$ is a minorant of $I + \Gamma(H, \jmath\omega)F_{\mathrm{npr}}(H_2, \jmath\omega)$ for all $H \in \mathbf{H}$. Now (3.43) implies that $I - \hat{\Gamma}(\jmath\omega)\hat{F}_{\mathrm{npr}}(\jmath\omega)$ is an M-matrix. Hence, it follows from Lemma 2.2 that $I + \Gamma(H, \jmath\omega)F_{\mathrm{npr}}(H_2, \jmath\omega)$ is invertible for all $H \in \mathbf{H}$ and $\omega \in (0,\infty)$. Futhermore, since by Lemma 3.1 $\Gamma(H, \jmath\omega)$ is invertible it follows that

$$\det[\Phi^{-1}(H_1, \jmath\omega) + F(H_2, \jmath\omega)] = \det[I + \Gamma(H, \jmath\omega)F_{\mathrm{npr}}(H_2, \jmath\omega)]\det[\Gamma^{-1}(H_1, \jmath\omega)].$$

Thus $\det[\Phi^{-1}(H_1, \jmath\omega) + F(H_2, \jmath\omega)] \neq 0$ for all $\omega \in (0,\infty)$. Now the performance bound (3.44) is a direct consequence of (2.2), (2.6), and (3.18c). $\qquad\square$

Remark 3.1. Note that if $F(H_2, s)$ is positive real then the spectral radius condition (3.43) is always satisfied, since $F_{\mathrm{npr}}(H_2, \jmath\omega) = 0$ for all $\omega \in (0,\infty)$. Hence Theorem 3.1 predicts unconditional stability for all uncertain positive real plants controlled by strictly

positive real compensators. Furthermore, in this case, the performance bound given by (3.44) collapses to the performance bound obtained in Hyland *et al.* (1994).

It is important to note that the results presented in this section are not restricted to positive real plants and positive real compensators. Specifically, if we assume that $\Phi^{-1}(H_1, s)$ is not positive real in (3.26) and define $G_{\theta w, \text{pr}}(H, \jmath\omega)$ and $G_{\theta w, \text{npr}}(H, \jmath\omega)$ to be the positive real and non-positive real parts, respectively, of $G_{\theta w}(H, \jmath\omega) \triangleq \Phi^{-1}(H_1, \jmath\omega) + F(H_2, \jmath\omega)$ such that

$$\lambda_{\min}(G_{\theta w, \text{pr}}(H, \jmath\omega) + G_{\theta w, \text{pr}}^{\text{H}}(H, \jmath\omega)) > 0$$

and

$$G_{\theta w, \text{npr}}(H, \jmath\omega) \triangleq G_{\theta w}(H, \jmath\omega) - G_{\theta w, \text{pr}}(H, \jmath\omega),$$

for all $\omega \in (0, \infty)$ and $H \in \mathbf{H}$, then Theorem 3.1 holds with minor modifications. Note that, in this case, no assumption on either the plant or the compensator is required.

4. Illustrative Numerical Example

In order to demonstrate the effectiveness of the proposed approach we present an illustrative example. Specifically, consider the simply supported Euler-Bernoulli beam with governing partial differential equation for the transverse deflection $w(x, t)$ given by

$$m(x)\frac{\partial^2 w(x, t)}{\partial t^2} + \frac{\partial^2}{\partial x^2}[EI(x)\frac{\partial^2 w(x, t)}{\partial x^2}] = f(x, t),$$

and with boundary conditions

$$w(x, t)|_{x=0, L} = 0, \quad EI(x)\frac{\partial^2 w(x, t)}{\partial x^2}|_{x=0, L} = 0,$$

where $m(x)$ is mass per unit length and $EI(x)$ is the flexural rigidity with E denoting Young's modulus of elasticity and $I(x)$ denoting the cross-sectional area moment of inertia about an axis normal to the plane of vibration and passing through the center of the cross-sectional area. Finally, $f(x, t)$ is the force distribution due to control actuation. Assuming uniform beam properties, the modal decomposition of this system has the form

$$w(x, t) = \sum_{r=1}^{\infty} W_r(x)q_r(t),$$

11

$$\int_0^L mW_r^2(x)dx = 1, \quad W_r(x) = \sqrt{\frac{2}{mL}}\sin\frac{r\pi x}{L}, \quad r = 1, 2, \ldots,$$

where, assuming uniform proportional damping, the modal coordinates q_r satisfy

$$\ddot{q}_r(t) + 2\zeta\Omega_r\dot{q}_r(t) + \Omega_r^2 q_r(t) = \int_0^L f(x,t)W_r(x)dx, \quad r = 1, 2, \ldots.$$

For simplicity assume $L = \pi$ and $m = EI = 2/\pi$ so that $\sqrt{\frac{2}{mL}} = 1$. Furthermore, we place a colocated velocity/force actuator pair at $x = 0.55L$. Finally, modeling the first two modes and defining the plant state as $x = [q_1\ \dot{q}_1\ q_2\ \dot{q}_2]^T$, and defining the performance of the beam in terms of the velocity at $x = 0.7L$, the resulting state space model and problem data are

$$\dot{x}(t) = \hat{A}x(t) + \hat{B}u(t) + D_1 w(t),$$

$$y(t) = \hat{C}x(t) + D_2 w(t),$$

where

$$\hat{A} = \text{block--diag}_{i=1,2}\begin{bmatrix} 0 & 1 \\ -\Omega_i^2 & -2\zeta\Omega_i \end{bmatrix}, \quad \Omega_i = i^2, \quad \zeta = 0.01,$$

$$\hat{B} = \hat{C}^T = [0\quad 0.9877\quad 0\quad -0.309]^T, \quad D_1 = [\hat{B}\ 0_{4\times 1}], \quad D_2 = [0\ 1.9],$$

with the performance variables

$$z(t) = E_1 x(t) + E_2 u(t),$$

where

$$E_1 = \begin{bmatrix} 0 & 0.809 & 0 & -0.951 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad E_2 = [0\quad 1.9]^T.$$

Using Theorem 3.2 of Haddad *et al.* (1994) we design a strictly positive real compensator $K(s)$. Next we assume frequency uncertainty in both Ω_1 and Ω_2 with $\widehat{\Delta\Omega}_1 = 0.5$ and $\widehat{\Delta\Omega}_2 = 0.4$. To reflect a more realistic setting, we include actuator and sensor dynamics described by

$$\Psi_a(s) = \frac{20}{s + 20}, \qquad \Psi_s(s) = \frac{20}{s + 20}.$$

Because of the actuator and sensor dynamics, $\Psi_a(s)K(s)\Psi_s(s)$ is positive real only up to $\omega = 2.5$ rad/sec as seen in Figure 1. Hence the techniques developed in Hyland *et*

12

al. (1994) for generating frequency domain performance bounds cannot be applied here. For the assumed uncertainty range the complex structured singular value bound (μ bound) (Hyland *et al.*, 1994; Packard and Doyle, 1993) and the complex block-structured majorant bound (Hyland and Collins, 1991) are infinite since both methods predict instability. The proposed majorant bound shown in Figure 2 gives a tight finite performance bound.

5. Conclusion

This paper developed frequency domain performance bounds for closed-loop uncertain positive real vibrational systems controlled by strictly positive real compensators along with appended actuator and sensor dynamics. These results are developed by using properties of logarthmic norms in conjuction with majorant analysis. The effectiveness of the proposed approach was demonstrated on a vibrational uncertain system with actuator and sensor dynamics.
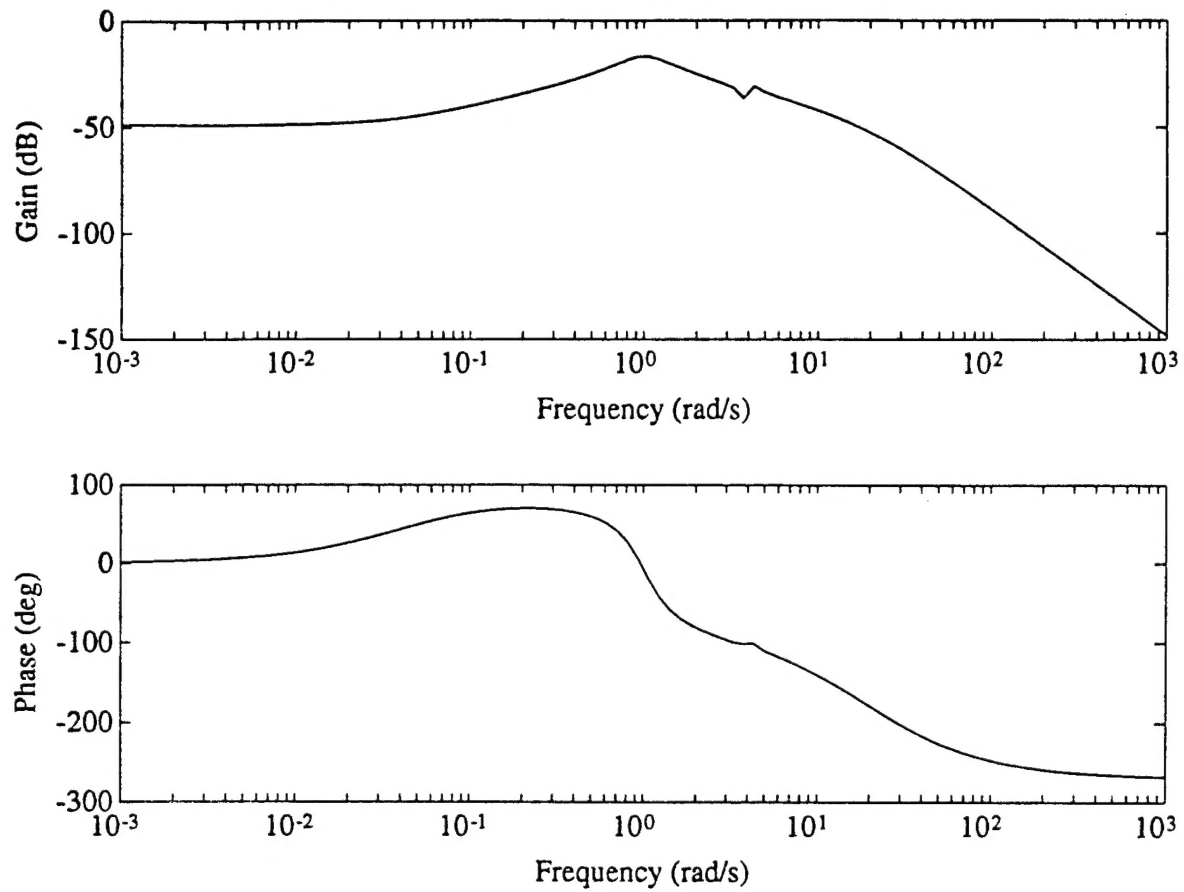
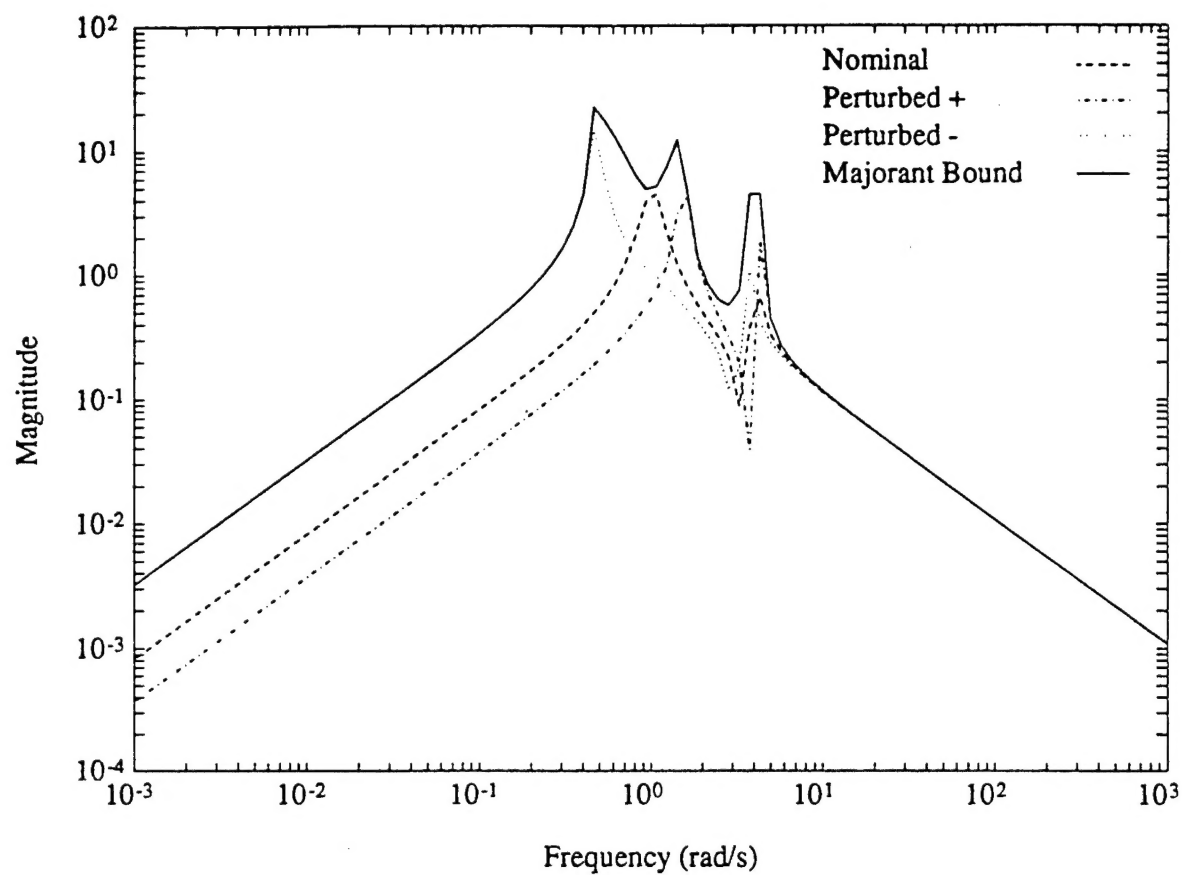**Figure 1.** Compensator with Actuator and Sensor Dynamics

Figure 2. Performance Bound for the Euler-Bernoulli Beam

References

Berman, A., and R.J. Plemmons (1979). *Nonnegative Matrices in the Mathematical Sciences*. Academic.

Bernstein, D.S., and W.M. Haddad (1990). Robust stability and performance analysis for state space systems via quadratic Lyapunov bounds. *SIAM J. Matrix Anal. Appl.*, **11**, 239-271.

Collins, Jr., E.G., and D.C. Hyland (1989). Improved robust performance bounds in covariance majorant analysis. *Int. J. Contr.*, **50**, 495-509.

Dahlquist, G. (1983). On matrix majorants and minorants with applications to differential equations. *Lin. Alg. Appl.*, **52/53**, 199-216.

Fiedler, M., and V. Ptak (1962). On matrices with non-positive off-diagonal elements and positive principal minors. *Czechoslovakian Math. J.*, **12**, 382-400.

Haddad, W.M., D.S. Bernstein, and Y.W. Wang. Dissipative H_2/H_∞ controller synthesis. *IEEE Trans. Autom. Contr.*, to appear.

Hyland, D.C., and D.S. Bernstein (1987). The majorant Lyapunov equation: A nonnegative matrix equation for guaranteed robust stability and performance of large scale systems. *IEEE Trans. Autom. Contr.*, **32**, 1005-1013.

Hyland, D.C., and E.G. Collins, Jr. (1989). An M-matrix and majorant approach to robust stability and performance analysis for systems with structured uncertainty. *IEEE Trans. Autom. Contr.*, **34**, 691-710.

Hyland, D.C., and E.G. Collins, Jr. (1991). Some majorant robustness results for discrete-time systems. *Automatica.*, **27**, 167-172.

Hyland, D.C., E.G. Collins, Jr., W.M. Haddad, V. Chellaboina (1994). Frequency domain performance bounds for uncertain positive real plants controlled by strictly positive real compensators. in *Proc. Amer. Contr. Conf.*

Ostrowski, A.M. (1975). On some metrical properties of operator matrices and matrices partitioned into blocks. *J. Math. Anal. Appl.*, **10**, 161-209.

Packard, A., and J.C. Doyle (1993). The complex structured singular value. *Automatica.*, **29**, 71-109.

Seneta, E. (1973). *Non-Negative Matrices*. Wiley.